

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 11:31:47 ; Search time 825 Seconds
(without alignments)
11003.891 Million cell updates/sec

Title: US-10-025-966A-4

Perfect score: 3363

Sequence: 1 gactccgcgcattcccaaga.....aaagaaaaa 3363

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3363	100.0	3363	24	ABSS56553
2	3356.8	99.8	3495	22	AAK52348
3	3348.4	99.6	4303	22	AAK05336
4	3348.4	99.6	4303	24	ABL90787
5	3347.2	99.5	4397	24	ABN86354
6	3346.8	99.5	4286	24	ABSS56559
7	3344.8	99.5	3348	22	AAH76201
8	3323.4	98.8	3906	21	AAZ64969

9	3323.4	98.8	3906	22	AAZ21253
10	3323.4	98.8	3906	22	AAF44115
11	3323.4	98.8	3906	24	ABL95597
12	3323.4	98.8	3906	24	ABL88108
13	3323.4	98.8	3906	25	ACR03612
14	3323.4	98.8	3906	25	ACR04033
15	3323.4	98.8	3906	25	ABX89150
16	3323.4	98.8	3906	25	ABX80171
17	3323.4	98.8	3906	25	ABX80675
18	3323.4	98.8	3906	25	ABX81058
19	3323.4	98.8	3906	25	ABX90148
20	3323.4	98.8	3906	25	ABX77759
21	3323.4	98.8	3906	25	ABX79355
22	3323.4	98.8	3906	25	ABX63994
23	3323.4	98.8	3906	25	ABX16958
24	3320.2	98.7	3335	22	AAK05304
25	3226.8	96.0	3435	22	AAK52349
26	3223.2	95.8	3513	24	ABN86352
27	3208.4	95.4	4137	21	AAK76477
28	2628	78.1	2901	22	AAK53333
29	2626.4	78.1	2901	22	AAK53332
30	2613	77.7	2616	24	ABSS56554
31	2605.2	77.5	2610	24	ABN86353
32	2294.8	68.2	2559	22	AAK94881
33	2235.2	66.5	2730	20	AAK84696
34	2174.8	64.7	2178	22	AAK94820
35	2095.2	62.3	3613	24	ABSS56561
36	2095.2	62.3	3766	24	ABSS56557
37	1073	31.9	5769	22	AAK03764
38	1040.4	30.9	2616	22	AAH46865
39	1040.4	30.9	4321	22	AAH46861
40	1040.4	30.9	4378	22	AAH72478
41	1040.4	30.9	4378	22	AAH72483
42	1040.4	30.9	4378	22	AAH72562
43	1040.4	30.9	4378	22	AAH72563
44	1040.4	30.9	4378	22	AAH72565
45	1040.4	30.9	4378	22	AAH72761

ALIGNMENTS

RESULT 1
ABSS56553
ID ABSS56553 standard; cDNA; 3363 BP.

XX ABSS56553;

AC ABSS56553;

DT 20-JAN-2003 (first entry)

XX cDNA encoding human SULF2 protein.

Human; SULF2; ss; gene; glucosamine-6-sulphatase; cancer; ischaemia;
tumour; angiogenesis; coronary; carotid; arterial occlusive disease;
peripheral arterial disease; atherosclerosis; myointimal hyperplasia;
thromboangitis obliterans; thrombotic disorder; vasculitis;
heart attack; myocardial infarction; vascular death; inflammation;
rheumatoid arthritis; asthma; adult respiratory distress syndrome;
sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;
allotraft rejection; lymphoma; thrombosis; sulphatase;
Chromosome 20q12-13.2.

XX Homo sapiens.

Key	Location/Qualifiers
FT sig_peptide	46..72
FT mat_peptide	73..2655
FT CDS	46..2658
FT	/*tag= a
FT	/product= "Human SULF2 protein"

Db 1381 TGTCAGGTGCTGAGTACAGACGGCGTGTGAGACGCTGGGACAGAGTGGCAGTGTGTG 1440
Qy 1441 GAGGACGCCACGGGGAAGCTGAAGCTGTCATTAAGTCAAGGGGCCCATGCGGCTGGCGCGC 1500
Db 1441 GAGGACGCCACGGGGAAGCTGAAGCTGTCATTAAGTCAAGGGGCCCATGCGGCTGGCGCGC 1500
Qy 1501 AGCAGAGCCCTCTCCAACTCTGTCGTCAGTACTAGCGGACGGGACGGGACGCGCTGAC 1560
Db 1501 AGCAGAGCCCTCTCCAACTCTGTCGTCAGTACTAGCGGACGGGACGGGACGCGCTGAC 1560
Qy 1561 TGTGACAGCGGGGACTACAGCTCAGCCCTGGCGGACGCGGAAAAAACTCTTCAAGAG 1620
Db 1561 TGTGACAGCGGGGACTACAGCTCAGCCCTGGCGGACGCGGAAAAAACTCTTCAAGAG 1620
Qy 1621 AAGTACAGGCGAGCTATGTCGCGAGTCGCTCCATCCGCTCAGTGGCCATCAGAGTGGAC 1680
Db 1621 AAGTACAGGCGAGCTATGTCGCGAGTCGCTCCATCCGCTCAGTGGCCATCAGAGTGGAC 1680
Qy 1681 GGCAGGGTGTACCAAGCTAGGCTGGGTGATGCGGCCAGCCGCAAACTCACCAGCGG 1740
Db 1681 GGCAGGGTGTACCAAGCTAGGCTGGGTGATGCGGCCAGCCGCAAACTCACCAGCGG 1740
Qy 1741 CACTGCGCAGGGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1800
Db 1741 CACTGCGCAGGGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1800
Qy 1801 GAGGGCTTCCCGACTACTCAGCGGCCAACCCCATTAAGTGCACATCGTGTACATC 1860
Db 1801 GAGGGCTTCCCGACTACTCAGCGGCCAACCCCATTAAGTGCACATCGTGTACATC 1860
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGGACCTGTACAGTCCCTGCGAGGCTGGAAA 1920
Db 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGGACCTGTACAGTCCCTGCGAGGCTGGAAA 1920
Qy 1921 GACCAAGCTGCATATGACACACAGATTAAGACCTGCGAGAACAAATTAAGAACCTG 1980
Db 1921 GACCAAGCTGCATATGACACACAGATTAAGACCTGCGAGAACAAATTAAGAACCTG 1980
Qy 1981 AGGGAAGTCCGAGGTCACTGAAGAAAAAGCGCCAGAGAGATGTGACTGTACAAAAATC 2040
Db 1981 AGGGAAGTCCGAGGTCACTGAAGAAAAAGCGCCAGAGAGATGTGACTGTACAAAAATC 2040
Qy 2041 AGTACACACCCAGCAAAAGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTC 2100
Db 2041 AGTACACACCCAGCAAAAGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTC 2100
Qy 2101 AGGAAGGCTGCAAGAGAGGACAGAGTGTGCTGTTCGGGAGCAGAGCGCAAGAG 2160
Db 2101 AGGAAGGCTGCAAGAGAGGACAGAGTGTGCTGTTCGGGAGCAGAGCGCAAGAG 2160
Qy 2161 AAACCTCGCAAGCTGCTCAAGCGCCTGCAGAACCAACGACACGTCGACGATGCGGCGCTC 2220
Db 2161 AAACCTCGCAAGCTGCTCAAGCGCCTGCAGAACCAACGACACGTCGACGATGCGGCGCTC 2220
Qy 2221 AGTGTCTTCCAC 2280
Db 2221 AGTGTCTTCCAC 2280
Qy 2281 TTCTGTGCTGCAACGCGCCAACTAATACAGTACTGTGTGATGAGGACCAATCAATGAG 2340
Db 2281 TTCTGTGCTGCAACGCGCCAACTAATACAGTACTGTGTGATGAGGACCAATCAATGAG 2340
Qy 2341 ACTCACAATTTCTCTCTGTGAATTTGCACTGGCTTCTTAGAGTACTTTGATCTCAAC 2400
Db 2341 ACTCACAATTTCTCTCTGTGAATTTGCACTGGCTTCTTAGAGTACTTTGATCTCAAC 2400
Qy 2401 ACAGACCCCTACAGCTGTATGATGAGTACAGTGAACACACTGACAGGAGTCTCTCAAC 2460
Db 2401 ACAGACCCCTACAGCTGTATGATGAGTACAGTGAACACACTGACAGGAGTCTCTCAAC 2460
Qy 2461 CTACAGCTACAGCTCTATGAGCTGAGGAGTCAAGGGTTTCAAGCAGTGAACCCCGG 2520
Db 2461 CTACAGCTACAGCTCTATGAGCTGAGGAGTCAAGGGTTTCAAGCAGTGAACCCCGG 2520

Qy 2521 ACTCGAAACATGACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGGCACTTT 2580
Db 2521 ACTCGAAACATGACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGGCACTTT 2580
Qy 2581 CAGGCTGAAAGTGGCCAGAAATGAAGAGACCTTCTTCAAATCACTGGGACAACTGTGG 2640
Db 2581 CAGGCTGAAAGTGGCCAGAAATGAAGAGACCTTCTTCAAATCACTGGGACAACTGTGG 2640
Qy 2641 GAAGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCCAAAAACATAGAGGATCAAC 2700
Db 2641 GAAGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCCAAAAACATAGAGGATCAAC 2700
Qy 2701 TGACTGCACAGGCAATGAATAAACAATGTGGTGAATTTCCAGCAGACGCTGTGCTATTGGCC 2760
Db 2701 TGACTGCACAGGCAATGAATAAACAATGTGGTGAATTTCCAGCAGACGCTGTGCTATTGGCC 2760
Qy 2761 AGGAGGCTCGAGAAAGCAACGACGCTCTCACTCACTCACTCACTCACTCACTCACTCACT 2820
Db 2761 AGGAGGCTCGAGAAAGCAACGACGCTCTCACTCACTCACTCACTCACTCACTCACTCACT 2820
Qy 2821 CAGCAGAGCAGAGATAAATTCAGGAAGTCCATTTTGGCCCTGCTTTTCTTTGGATTA 2880
Db 2821 CAGCAGAGCAGAGATAAATTCAGGAAGTCCATTTTGGCCCTGCTTTTCTTTGGATTA 2880
Qy 2881 TACCTCAGCAGCTGCACAAAATGCAATTTTCTGATCAAAAAGTCACTAACCTTCCCTCC 2940
Db 2881 TACCTCAGCAGCTGCACAAAATGCAATTTTCTGATCAAAAAGTCACTAACCTTCCCTCC 2940
Qy 2941 CCAGAGCTCACAAGGAAAACGAGAGAGCGAGAGAGAGATTTTCTTTGGAAATTTCT 3000
Db 2941 CCAGAGCTCACAAGGAAAACGAGAGAGCGAGAGAGAGATTTTCTTTGGAAATTTCT 3000
Qy 3001 CCCAAGGGGAAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTAAA 3060
Db 3001 CCCAAGGGGAAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTGTTCTAAA 3060
Qy 3061 TCCTCTTATTTCTTTGGTTGTCAAAAAGGAACTTAAGAGCAGACAGAGGCAACGT 3120
Db 3061 TCCTCTTATTTCTTTGGTTGTCAAAAAGGAACTTAAGAGCAGACAGAGGCAACGT 3120
Qy 3121 GGAGAGCTGAAAACAGTGCAGAGAGCTTTGCAATGAGTCACTAGCAGCAAAAAGAGATGA 3180
Db 3121 GGAGAGCTGAAAACAGTGCAGAGAGCTTTGCAATGAGTCACTAGCAGCAAAAAGAGATGA 3180
Qy 3181 CATTACCTAGCACTATAACCTGCTGCTGAGGAACTGCTTCACTGATATAT 3240
Db 3181 CATTACCTAGCACTATAACCTGCTGCTGAGGAACTGCTTCACTGATATAT 3240
Qy 3241 GTGACTATTTTACATGTAATCAACATGGAACTTTTAGGGGAACTTAAGAAATCCCAA 3300
Db 3241 GTGACTATTTTACATGTAATCAACATGGAACTTTTAGGGGAACTTAAGAAATCCCAA 3300
Qy 3301 TTTTCAGGAGTGGTGTGTCATTAACGCTCTGTGGCAGTGTAAAGAAAAA 3360
Db 3301 TTTTCAGGAGTGGTGTGTCATTAACGCTCTGTGGCAGTGTAAAGAAAAA 3360
Qy 3361 AAA 3363
Db 3361 AAA 3363

RESULT 2

AAK52348

ID AAK52348 standard; cDNA; 3495 BP.

XX AAK52348;

AC AAK52348;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 893.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200157190-A2.

FN 09-AUG-2001.

PD 05-FEB-2001; 2001WO-US04098.

PF 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

DR P-PSDB; AAM79215.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
 useful in diagnosis and gene therapy -

PT Claim 1; Page 2939-2943; 6221pp; English.

PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
 encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX Sequence 3495 BP; 958 A; 976 C; 924 G; 637 T; 0 other;

Query Match 99.8%; Score 3356.8; DB 22; Length 3495;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3358; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GACTCCGCGATCCCAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 60

Db 136 GACTCCGCGATCCCAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 195

Qy 61 CTCGTGCTGTGCTGTGCTGCGCACTGTGTTCTCCCTGCTGGTGAAGCTGGCTTC 120

Db 196 CTCGTGCTGTGCTGTGCTGCGCACTGTGTTCTCCCTGCTGGTGAAGCTGGCTTC 255

Qy 121 CTGTGCGACACACCGCTGAAAGGAGGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 180

Db 256 CTGTGCGACACACCGCTGAAAGGAGGTTTCAGAGGACCGCAGGAACATCCGCCCAAC 315

Qy 181 ATCATCTGTGTGTCAGGACGACACAGATGTGGAGTGGGTTCATGACAGGTGATGAAC 240

Db 316 ATCATCTGTGTGTCAGGACGACACAGATGTGGAGTGGGTTCATGACAGGTGATGAAC 375

Qy 241 AAGACCGGCGCATCATGGACGAGGGGGCGGCACATTCATCAACGCTTCGTGACCA 300

Db 376 AAGACCGGCGCATCATGGACGAGGGGGCGGCACATTCATCAACGCTTCGTGACCA 435

Qy 301 CCCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

Db 436 CCCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495

Qy 361 ACCTACACCAACAAATGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 496 ACCTACACCAACAAATGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555

Qy 421 ACCTTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Db 556 ACCTTTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615

Qy 481 AATGAATAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Db 616 AATGAATAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675

Qy 541 AACTCCCGCTTTTATAACTACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Db 676 AACTCCCGCTTTTATAACTACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735

Qy 601 GACTACTCCAAAGGATTAACCTCAGAGACCTCATACCAATGACAGCGTGAGCTTCTTCGCG 660

Db 736 GACTACTCCAAAGGATTAACCTCAGAGACCTCATACCAATGACAGCGTGAGCTTCTTCGCG 795

Qy 661 ACCTCCAAAGAGATGTACCCGCGACAGGCCAGTCTCATGTGCTCATGCCATGACGCCCTCC 720

Db 796 ACCTCCAAAGAGATGTACCCGCGACAGGCCAGTCTCATGTGCTCATGCCATGACGCCCTCC 855

Qy 721 CACGCGCTGAGGATTCAGCCCGACCAATATTCAGCGCTCTTCCCAACGCGATCTCAGCAC 780

Db 856 CACGCGCTGAGGATTCAGCCCGACCAATATTCAGCGCTCTTCCCAACGCGATCTCAGCAC 915

Qy 781 ATCAGCGCGAGCTCAACTACGCGCCCAACCCCGGACAAACACTGATCATCGCTGCTACAG 840

Db 916 ATCAGCGCGAGCTCAACTACGCGCCCAACCCCGGACAAACACTGATCATCGCTGCTACAG 975

Qy 841 GGGCCCATGAAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAGAGCGCTTCAG 900

Db 976 GGGCCCATGAAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAGAGCGCTTCAG 1035

Qy 901 ACCCTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

Db 1036 ACCCTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1095

Qy 961 GAGCTGGAACACAGTACATCGTATACACCGCGACACCGGTTACCAATCGCGCGAGTTT 1020

Db 1096 GAGCTGGAACACAGTACATCGTATACACCGCGACACCGGTTACCAATCGCGCGAGTTT 1155

Qy 1021 GGCCTGTGTAAGGGAATTCATGCCATATGAGTTTGACATCAGGCTCCCGTTCCTAGCTG 1080

Db 1156 GGCCTGTGTAAGGGAATTCATGCCATATGAGTTTGACATCAGGCTCCCGTTCCTAGCTG 1215

Qy 1081 AGGGGCCCCAACGCTGGAAGCGGCTGTCTGAATCCCACTGCTCTCAACATTTGACCTG 1140

Db 1216 AGGGGCCCCAACGCTGGAAGCGGCTGTCTGAATCCCACTGCTCTCAACATTTGACCTG 1275

Qy 1141 GCGCCACCATCTGTCGACATTTGAGGCTGACATACCTCGGATATGGAGGGAATCC 1200

Db 1276 GCGCCACCATCTGTCGACATTTGAGGCTGACATACCTCGGATATGGAGGGAATCC 1335

Qy 1201 ATCTCTCAAGCTGCTGGACACGAGCGCGCTGATGATCGGTTTCACTTGAAGAAGATG 1260

Db 1336 ATCTCTCAAGCTGCTGGACACGAGCGCGCTGATGATCGGTTTCACTTGAAGAAGATG 1395

Qy 1261 AGGTCTGCGGAGTCTGCTTCTGCTGAGAGAGGCAAGCTGTCTACACAGAGAGACAT 1320

Db 1396 AGGTCTGCGGAGTCTGCTTCTGCTGAGAGAGGCAAGCTGTCTACACAGAGAGACAT 1455

Qy 1321 GACAGGTGAGCGCGCCAGGAGGAACTTTCTGCGCCCAAGTACACAGCGTGTGAAGGACCTG 1380

Db 1456 GACAAGGTGGACCCGAGGAGAGAACTTTCTGCCCAAGTACACAGCGTGTGAAGACCTG 1515
Qy 1381 TGTCAAGCGTCTCAGTACACAGACGGCGTGTGACACCTGGGACAGAAAGTGGCAGTGTGTG 1440
Db 1516 TGTCAAGCGTCTCAGTACACAGACGGCGTGTGACACCTGGGACAGAAAGTGGCAGTGTGTG 1575
Qy 1441 GAGGACGCCACCGGGAGCTGAAGCTGCATAGTGAAGGGGCCCCATGCGGTGGGCGGC 1500
Db 1576 GAGGACGCCACCGGGAGCTGAAGCTGCATAGTGAAGGGGCCCCATGCGGTGGGCGGC 1635
Qy 1501 AGCAGAGCCCTCTCCAACTCTGCGCCCAAGTACTACGGGACGGGACGAGCCCTGCAC 1560
Db 1636 AGCAGAGCCCTCTCCAACTCTGCGCCCAAGTACTACGGGACGGGACGAGCCCTGCAC 1695
Qy 1561 TGTGACAGCGGGACACTCAAGCTCAGCCTGGCCGAGCGCCGGAAGAACTCTTCAAGAG 1620
Db 1696 TGTGACAGCGGGACACTCAAGCTCAGCCTGGCCGAGCGCCGGAAGAACTCTTCAAGAG 1755
Qy 1621 AAGTACAGGCGCACTATGTCCGAGTCTGCTCCATCGCTCAGTGGCCATCAGGTGAC 1680
Db 1756 AAGTACAGGCGCACTATGTCCGAGTCTGCTCCATCGCTCAGTGGCCATCAGGTGAC 1815
Qy 1681 GGCAGAGGTGTACCACTGAGGCTTGGGTGATGCCGCCAGCCCGGAACTCCACCAAGCG 1740
Db 1816 GGCAGAGGTGTACCACTGAGGCTTGGGTGATGCCGCCAGCCCGGAACTCCACCAAGCG 1875
Qy 1741 CACTGGCCAGGGGCCCTGAGGACCAAGATGACAAGATGTGGGACCTTCAAGTGGCACT 1800
Db 1876 CACTGGCCAGGGGCCCTGAGGACCAAGATGACAAGATGTGGGACCTTCAAGTGGCACT 1935
Qy 1801 GGAGGCTTCCGACACTACTCAGCGCGCCCAACCCATTAAAGTGCACATCGGTGTACATC 1860
Db 1936 GGAGGCTTCCGACACTACTCAGCGCGCCCAACCCATTAAAGTGCACATCGGTGTACATC 1995
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGTGACAGTCCCTGAGGCTTGGAAA 1920
Db 1996 CTAGAGAACGACACAGTCCAGTGTGACCTGTGACAGTCCCTGAGGCTTGGAAA 2055
Qy 1921 GACCAAGCTGCACATCGACACGAGATGAACCTCGCAGCAAAATTAAGACCTG 1980
Db 2056 GACCAAGCTGCACATCGACACGAGATGAACCTCGCAGCAAAATTAAGACCTG 2115
Qy 1981 AGGGAAGTCCGAGGTCACTTGAAGAAAAAGCGGCCAGAAAGTGTGACTGTCACAAAATC 2040
Db 2116 AGGGAAGTCCGAGGTCACTTGAAGAAAAAGCGGCCAGAAAGTGTGACTGTCACAAAATC 2175
Qy 2041 AGCTACACACCCAGCAAAAGCGCGCTCAAGCAAGAGGCTTCCAGTGTGATCTTTC 2100
Db 2176 AGCTACACACCCAGCAAAAGCGCGCTCAAGCAAGAGGCTTCCAGTGTGATCTTTC 2235
Qy 2101 AGGAAGGCTTCAAGAGAGGCAAGGTGTGGCTGTGGGAGCAGAAAGCGCAAGAG 2160
Db 2236 AGGAAGGCTTCAAGAGAGGCAAGGTGTGGCTGTGGGAGCAGAAAGCGCAAGAG 2295
Qy 2161 AAATCCGCAAGTGTCTCAAGCGCTTCAAGCAAAAGCGCTGACAGATGCGAGCGCTC 2220
Db 2296 AAATCCGCAAGTGTCTCAAGCGCTTCAAGCAAAAGCGCTGACAGATGCGAGCGCTC 2355
Qy 2221 ACCTGCTTCAACCAAGCAAGCGCTTCAAGCAAGCGCGCTTTCGACACTGGGCGCT 2280
Db 2356 ACCTGCTTCAACCAAGCAAGCGCTTCAAGCAAGCGCGCTTTCGACACTGGGCGCT 2415
Qy 2281 TTCTGTCTGCACAGCGCCCAACAAATCAACCTGCTGTGCTAGTGGACCATCAATGAG 2340
Db 2416 TTCTGTCTGCACAGCGCCCAACAAATCAACCTGCTGTGCTAGTGGACCATCAATGAG 2475
Qy 2341 ACTCAAAATTTCTCTCTGTGAATTTGAACTTGGCTTCTAGAGTACTTGTATCTCAAC 2400
Db 2476 ACTCAAAATTTCTCTCTGTGAATTTGAACTTGGCTTCTAGAGTACTTGTATCTCAAC 2535
Qy 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACTGACAGAGGATGTCTTCAACAG 2460
Db 2536 ACAGACCCCTACAGCTGATGATGAGTGAACACACTGACAGAGGATGTCTTCAACAG 2595

Qy 2461 CTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGG 2520
Db 2596 CTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGG 2655
Qy 2521 ACTCGAAACATGACCTGGGACCTTAAAGATGAGGAAAGCTATAGCAATACAGCAGTTT 2580
Db 2656 ACTCGAAACATGACCTGGGACCTTAAAGATGAGGAAAGCTATAGCAATACAGCAGTTT 2715
Qy 2581 CAGCGTCCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCAGCTGGGACAACTGTGG 2640
Db 2716 CAGCGTCCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCAGCTGGGACAACTGTGG 2775
Qy 2641 GAAGCTGGGAAGTTAAGAAACAACAGAGGTGACCTCCAAACATAGAGGATCAAC 2700
Db 2776 GAAGCTGGGAAGTTAAGAAACAACAGAGGTGACCTCCAAACATAGAGGATCAAC 2835
Qy 2701 TGACTGCACAGGCAATGAAACCAATGCTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 2836 TGACTGCACAGGCAATGAAACCAATGCTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2895
Qy 2761 AGGAGGCTCGAAGAACGACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 2896 AGGAGGCTCGAAGAACGACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2955
Qy 2821 CAGCAGAGCAGAGATAAATCTCAGGAAGTCAATTTTGGCCCTGCTTTGCTTTGGATTA 2880
Db 2956 CAGCAGAGCAGAGATAAATCTCAGGAAGTCAATTTTGGCCCTGCTTTGCTTTGGATTA 3015
Qy 2881 TACTCTCACCAGTCCACAAATGCAATTTTTCGTATCAAAAGTCAACACTTAACCCCTCCC 2940
Db 3016 TACTCTCACCAGTCCACAAATGCAATTTTTCGTATCAAAAGTCAACACTTAACCCCTCCC 3075
Qy 2941 CCAGAAGCTCAAAAGGAAACGGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3076 CCAGAAGCTCAAAAGGAAACGGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3135
Qy 3001 CCCAAGGCGCAAGTCAATGGAATTTTAAATCATAGGGAAAGCAGTCTGTCTTAAA 3060
Db 3136 CCCAAGGCGCAAGTCAATGGAATTTTAAATCATAGGGAAAGCAGTCTGTCTTAAA 3195
Qy 3061 TCCTCTTATTCTTTTGGTTTGTCAAAAGAGGAATTAAGAGCAGGACGAGGCAACGT 3120
Db 3196 TCCTCTTATTCTTTTGGTTTGTCAAAAGAGGAATTAAGAGCAGGACGAGGCAACGT 3255
Qy 3121 GGAGAGCTGAAACAGTGCAGAGACCTTTGCAATGAGTCACTAGCACAAGAGATGA 3180
Db 3256 GGAGAGCTGAAACAGTGCAGAGACCTTTGCAATGAGTCACTAGCACAAGAGATGA 3315
Qy 3181 CATTTACTAGCACTATAAAACCCCTGGTTGCTCTGAAAGAACTGCCCTTCATTGTATAT 3240
Db 3316 CATTTACTAGCACTATAAAACCCCTGGTTGCTCTGAAAGAACTGCCCTTCATTGTATAT 3375
Qy 3241 GTGACTATTACATGTAATCAAGTGGAACTTTTAGGGAACTTAATAAGAAATCCCAA 3300
Db 3376 GTGACTATTACATGTAATCAAGTGGAACTTTTAGGGAACTTAATAAGAAATCCCAA 3435
Qy 3301 TTTTCAGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAAAGAAAAA 3360
Db 3436 TTTTCAGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAAAGAAAAA 3495

RESULT 3

AAD05336
ID AAD05336 standard; cDNA; 4303 BP.

XX AAD05336;

XX 17-JUL-2001 (first entry)

DE Human secreted protein-encoding gene 5 cDNA clone HE9QN39, SEQ ID NO:47.

KW Human; secreted protein; proliferative disorder; cancer; tumour;

foetal abnormality; developmental abnormality; haematopoietic disorder;
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
inflammation; allergy; neurological disorder; Alzheimer's disease;
Parkinson's disease; cognitive disorder; schizophrenia; asthma;
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
cardiovascular disorder; angiotensin disorder; kidney disorder;
gastrointestinal disorder; pregnancy-related disorder;
endocrine disorder; infection; wound healing; vulvectomy;
cell culture; chemotaxis; food additive; gene therapy;
binding partner identification; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 222..2834

FT /product= "Human secreted protein"
FT /transl_except= [pos:1821..1823, aa:Xaa]
FT /note= "Xaa equals any of the twenty naturally occurring
FT L-amino acids"
FT /tag= a
FT /tag= b
FT /tag= c
FT /product= "Mature human secreted protein"

XX WO200134626-A1.

XX 17-MAY-2001.

XX 01-NOV-2000; 2000WO-US30045.

XX 05-NOV-1999; 99US-0163581.

XX 30-JUN-2000; 2000US-021513.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;

XX WPI: 2001-308778/32.

XX P-PSDB; AAE01471.

XX New nucleic acid molecules encoding 28 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives -

XX Claim 1; Page 450-452; 562pp; English.

XX AAD05300-AA05379 represent cDNAs corresponding to 28 human secreted
XX protein genes, and AAE01436-AAE01513 represent the proteins they encode.
XX AAE01514-AAE01544 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 28 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in

CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein-encoding cDNA of the invention.

XX Sequence 4303 BP; 1195 A; 1142 C; 1073 G; 889 T; 4 other;

XX Query Match 99.6%; Score 3348.4; DB 22; Length 4303;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 3349; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACTCCCGCATCCCAAGAAGACACAGATCAGCAAAAAGAGATGGGCCCCCGGAGC 60
DB 177 GACTCCCGCATCCCAAGAAGACACAGATCAGCAAAAAGAGATGGGCCCCCGGAGC 236
QY 61 CTCGTGCTGTGCTTCTGCTCGCAACTGTGTCTCCCTGCTGGGTGGAGTCTGGGCTTC 120
DB 237 CTCGTGCTGTGCTTCTGCTCGCAACTGTGTCTCCCTGCTGGGTGGAGTCTGGGCTTC 296
QY 121 CTGTGCGACCAACCGCTGAAAGGCGAGTTTCAGAGGACCGCAGGAAACATCCGCCCAAC 180
DB 297 CTGTGCGACCAACCGCTGAAAGGCGAGTTTCAGAGGACCGCAGGAAACATCCGCCCAAC 356
QY 181 ATCATCTCTGCTGTGACGAGACGACGAGATGTGGAGTGGGTTCATCAGTGTATGAAC 240
DB 357 ATCATCTCTGCTGTGACGAGACGACGAGATGTGGAGTGGGTTCATCAGTGTATGAAC 416
QY 241 AAGACCCCGCGCATCATGAGAGCGGGCGGCGCACTTCATCAAGCGCTTCGTGACCAACA 300
DB 417 AAGACCCCGCGCATCATGAGAGCGGGCGGCGCACTTCATCAAGCGCTTCGTGACCAACA 476
QY 301 CCCATGTCTGCCCTCACGCTCCATCTCTCACTGCGCAAGTACGTCCACAACCAAC 360
DB 477 CCCATGTCTGCCCTCACGCTCCATCTCTCACTGCGCAAGTACGTCCACAACCAAC 536
QY 361 ACCTACACCAACAATGAGAACTGCTCTCGCCCTCTCTGGCAGGCACAGCAGAGAGCGGC 420
DB 537 ACCTACACCAACAATGAGAACTGCTCTCGCCCTCTCTGGCAGGCACAGCAGAGAGCGGC 596
QY 421 ACCTTGGCGGTACTCAATAGCACTGGCTACCGGACAGCTTTCTCGGGAAGTATCTT 480
DB 597 ACCTTGGCGGTACTCAATAGCACTGGCTACCGGACAGCTTTCTCGGGAAGTATCTT 656
QY 481 AATGAATACAAACGGCTCTCTACGTCACCCCGCTCGGAGAGTGGGTGCGACTCTTAAA 540
DB 657 AATGAATACAAACGGCTCTCTACGTCACCCCGCTCGGAGAGTGGGTGCGACTCTTAAA 716
QY 541 AACTCCCGCTTTTAACTAATACAGCTGTGTGGAAGCGGTGAAAGAAAGACAGCGCTCC 600
DB 717 AACTCCCGCTTTTAACTAATACAGCTGTGTGGAAGCGGTGAAAGAAAGACAGCGCTCC 776
QY 601 GACTACTCCAAGGATTACTCTACAGACTCTATCAATGACAGCTTCTTCGCGC 660
DB 777 GACTACTCCAAGGATTACTCTACAGACTCTATCAATGACAGCTTCTTCGCGC 836
QY 661 AGCTCAAGAAAGATGATACCGCAGAGCGAGTCTCTATGGTTCATCAGCATGACGCCCTC 720
DB 837 AGCTCAAGAAAGATGATACCGCAGAGCGAGTCTCTATGGTTCATCAGCATGACGCCCTC 896
QY 721 CAGCGCCCTGAGGTTACGCCCCCAATATTCACGCTCTTCCCAAGCGCATCTCAGCAC 780
DB 897 CAGCGCCCTGAGGTTACGCCCCCAATATTCACGCTCTTCCCAAGCGCATCTCAGCAC 956
QY 781 ATCAGCCCGAGTCAACAATCAGCGCCCAACCGGCAAACTGGATCATGCGTACACG 840
DB 957 ATCAGCCCGAGTCAACAATCAGCGCCCAACCGGCAAACTGGATCATGCGTACACG 1016
QY 841 GGGCCCATGAAGCCCATCCCATGGAATTCACCAATGCTCCAGCGGAAGCGTTGAG 900
DB 1017 GGGCCCATGAAGCCCATCCCATGGAATTCACCAATGCTCCAGCGGAAGCGTTGAG 1076
QY 901 ACCTCATCTCGGTGGAGCGACTTCCATGGAGAGGATTTACAAATGCTGTTGAGACGGGC 960

1077 ACCCTCATGTCGGTGGAGCGACTCCATGAGAGAGATTATCAACATGCTGGTTGAGACGGCG 1136
 961 GAGCTGGCAACACAGCTATACATCGTATATACACCGCGACCAAGCTTACCAATCGGCGAGTTT 1020
 1137 GAGCTGGCAACACAGCTATACATCGTATATACACCGCGACCAAGCTTACCAATCGGCGAGTTT 1196
 1021 GGCCTGGTGAAGGGAATCCATGCGCATATGAGTTTGAATCAGAGGTCCTGTTTACGTTG 1080
 1197 GGCCTGGTGAAGGGAATCCATGCGCATATGAGTTTGAATCAGAGGTCCTGTTTACGTTG 1256
 1081 AGGGGCCCCAAGCTGGAGCGGGCTGTGTAATCCCACTGCTCCTCAACATTTGACCTG 1140
 1257 AGGGGCCCCAAGCTGGAGCGGGCTGTGTAATCCCACTGCTCCTCAACATTTGACCTG 1316
 1141 GCGCCCAACATCCTCGGACATTTGAGGCTGGACATACCTCGGATATGAGCGGAATTC 1200
 1317 GCGCCCAACATCCTCGGACATTTGAGGCTGGACATACCTCGGATATGAGCGGAATTC 1376
 1201 ATCCTCAAGCTCTGGACACGAGCGGGCTGTAATCGGTTTCACTTGAAGAAAGAGATG 1260
 1377 ATCCTCAAGCTCTGGACACGAGCGGGCTGTAATCGGTTTCACTTGAAGAAAGAGATG 1436
 1261 AGGGTCTGGCGGACTCCTTCTTGGTGGAGAGCAAGCTGCTACACAGAGAGACAAT 1320
 1437 AGGGTCTGGCGGACTCCTTCTTGGTGGAGAGCAAGCTGCTACACAGAGAGACAAT 1496
 1321 GACAGGTGGACCGCCAGGAGAGAACTTCTGCGCAAGTACCAGCTGTGAAGACCTG 1380
 1497 GACAGGTGGACCGCCAGGAGAGAACTTCTGCGCAAGTACCAGCTGTGAAGACCTG 1556
 1381 TGTCAAGCTGTGAGTACACAGACGCGTGTGAGCAGCTGGACAGAACTGGCAGTGTGTG 1440
 1557 TGTCAAGCTGTGAGTACACAGACGCGTGTGAGCAGCTGGACAGAACTGGCAGTGTGTG 1616
 1441 GAGGACGCGCACGGGGAAGCTGAAGCTGCATAGTGCAGAGGCGCCCATCGGCTGGCGGCG 1500
 1617 GAGGACGCGCACGGGGAAGCTGAAGCTGCATAGTGCAGAGGCGCCCATCGGCTGGCGGCG 1676
 1501 AGCAGAGCCCTCTCAACCTCGTGGCCCAAGTACTACGGGCGAGGCGAGGCGCTGACCC 1560
 1677 AGCAGAGCCCTCTCAACCTCGTGGCCCAAGTACTACGGGCGAGGCGAGGCGCTGACCC 1736
 1561 TGTGACAGCGGGAATACAGCTCAGCTGGCGGACCGGACCGGAAAAAACTCTTCAAGAG 1620
 1737 TGTGACAGCGGGAATACAGCTCAGCTGGCGGACCGGACCGGAAAAAACTCTTCAAGAG 1796
 1621 AAGTACAGGCGAGTATGTCGAGTGCCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1680
 1797 AAGTACAGGCGAGTATGTCGAGTGCCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1856
 1681 GGCAGGGTGTACACAGTAGGCTGGGTGATGCGGCCCGGACCGGAAAACTCACCAGCGG 1740
 1857 GGCAGGGTGTACACAGTAGGCTGGGTGATGCGGCCCGGACCGGAAAACTCACCAGCGG 1916
 1741 CACTGGCAGGGGCCCCGTGAGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1800
 1917 CACTGGCAGGGGCCCCGTGAGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1976
 1801 GGAGGCTTCCCGACTACTCAGCGCGCAACCCATTAAGTGAACATCGGTGTACATC 1860
 1977 GGAGGCTTCCCGACTACTCAGCGCGCAACCCATTAAGTGAACATCGGTGTACATC 2036
 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTCAGGCTGGAAA 1920
 2037 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTCAGGCTGGAAA 2096
 1921 GACCAAGCTGCATCGACACGAGTTGAACCCCTGCGAACAAATTAAGAACCTG 1980
 2097 GACCAAGCTGCATCGACACGAGTTGAACCCCTGCGAACAAATTAAGAACCTG 2156
 1981 AGGGAAGTCCGAGGTCACTGAAAGAAAAAGCGCGCGCAGAGATGTGACTGTCACAAAATC 2040
 2157 AGGGAAGTCCGAGGTCACTGAAAGAAAAAGCGCGCGCAGAGATGTGACTGTCACAAAATC 2216

2041 AGCTACCACACCAGCAAAAGCGCGCTCAAGCAACAGAGGCTCCAGTGTGATCCTTTTC 2100
 2217 AGCTACCACACCAGCAAAAGCGCGCTCAAGCAACAGAGGCTCCAGTGTGATCCTTTTC 2276
 2101 AGGAAGGGCTGACAGAGAAAGAGCAAGGTGTGGCTGTGCGGAGCAGAAAGCGCAAGAG 2160
 2277 AGGAAGGGCTGACAGAGAAAGAGCAAGGTGTGGCTGTGCGGAGCAGAAAGCGCAAGAG 2236
 2161 AAACCTCCGAAGCTGCTCAAGCGCTGACAGAAACACACACGCTGACGATGCCAGGCTTC 2220
 2337 AAACCTCCGAAGCTGCTCAAGCGCTGACAGAAACACACACGCTGACGATGCCAGGCTTC 2396
 2221 ACCTGCTTCAACCAACAGCACTGGCAGACGCGCGCTTTCTGGACATCTGGGCGCT 2280
 2397 ACCTGCTTCAACCAACAGCACTGGCAGACGCGCGCTTTCTGGACATCTGGGCGCT 2456
 2281 TTCTGTGCTGACAGCGCCCAACATTAACATGCTGCTGATGAGGACCATCAATGAG 2340
 2457 TTCTGTGCTGACAGCGCCCAACATTAACATGCTGCTGATGAGGACCATCAATGAG 2516
 2341 ACTCAAAATTTCTTCTGCTGTAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2400
 2517 ACTCAAAATTTCTTCTGCTGTAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2576
 2401 ACAGACCCCTACAGCTGATGAGTGAACACACTGGACAGGATGCTCTCAACAG 2460
 2577 ACAGACCCCTACAGCTGATGAGTGAACACACTGGACAGGATGCTCTCAACAG 2636
 2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGAACCCCGG 2520
 2637 CTACAGCTACAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGAACCCCGG 2696
 2521 ACTCGAAAACATGACCTGGACTTAAAGATGAGAGAGCTATGAGCAATACAGGAGTTT 2580
 2697 ACTCGAAAACATGACCTGGACTTAAAGATGAGAGAGCTATGAGCAATACAGGAGTTT 2756
 2581 CAGCGTCGAAAGTGGCCGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
 2757 CAGCGTCGAAAGTGGCCGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2816
 2641 GAAGGTGGGAAGGTTAAGAAACACAGAGGTGGACCTTCAAAAAACATAGAGGATCAAC 2700
 2817 GAAGGTGGGAAGGTTAAGAAACACAGAGGTGGACCTTCAAAAAACATAGAGGATCAAC 2876
 2701 TGACTCAGACGCAATGAAGAACCACTGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760
 2877 TGACTCAGACGCAATGAAGAACCACTGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2936
 2761 AGGAGCCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTTGGAGGATAAC 2820
 2937 AGGAGCCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTTGGAGGATAAC 2996
 2821 CAGCAGGACAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880
 2997 CAGCAGGACAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3056
 2881 TACTCACCAGCTGCAAAAAATGCAATTTTTCGTATCAAAAAAGTCAACCACTAACCTCCC 2940
 3057 TACTCACCAGCTGCAAAAAATGCAATTTTTCGTATCAAAAAAGTCAACCACTAACCTCCC 3116
 2941 CAGAGAGCTCAAAAGGAAAACGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3000
 3117 CAGAGAGCTCAAAAGGAAAACGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3176
 3001 CCCAAGGGCGAAAGTCAATTTGGAAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAAA 3060
 3177 CCCAAGGGCGAAAGTCAATTTGGAAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAAA 3236
 3061 TCCTCTTATTTCTTTGGTTGTCAAAAGAGAACTAAGAGAGGAGCAGAGGCAACCT 3120
 3237 TCCTCTTATTTCTTTGGTTGTCAAAAGAGAACTAAGAGAGGAGCAGAGGCAACCT 3296

QY 961 GAGCTGGACAAACGTAACGTAATACCGCGGACCAAGGTTTACCAATCGCGCAGTTT 1020
DB 1137 GAGCTGGACAAACGTAACGTAATACCGCGGACCAAGGTTTACCAATCGCGCAGTTT 1196
QY 1021 GGCTGGTGAAGGGAATCCATGCAATGATGATTTGACATCAGGGTCCGTTTACGGT 1080
DB 1197 GGCTGGTGAAGGGAATCCATGCAATGATGATTTGACATCAGGGTCCGTTTACGGT 1256
QY 1081 AGGGGCCCAACGTTGAAGCGGCTGTCTGAATCCCAATCGCTCCCAATGACCTG 1140
DB 1257 AGGGGCCCAACGTTGAAGCGGCTGTCTGAATCCCAATCGCTCCCAATGACCTG 1316
QY 1141 GCGCCCAACATCTCGACATGCAAGGCTGGAATACCTGCGGATGATGACGCGGAATCC 1200
DB 1317 GCGCCCAACATCTCGACATGCAAGGCTGGAATACCTGCGGATGATGACGCGGAATCC 1376
QY 1201 ATCTCAAGCTGTGGACACGAGCGGCGGTGAATCGGTTTCACTTGAAGAAGAGATG 1260
DB 1377 ATCTCAAGCTGTGGACACGAGCGGCGGTGAATCGGTTTCACTTGAAGAAGAGATG 1436
QY 1261 AGGCTGTGGCGGACCTCTTCTTGTGGAGAGGCAAGCTGCTACCAAGAGAGCAAT 1320
DB 1437 AGGCTGTGGCGGACCTCTTCTTGTGGAGAGGCAAGCTGCTACCAAGAGAGCAAT 1496
QY 1321 GACAAGGTGACCGCCAGGAGGAGAACTTCTGCCCAAGTACCAAGGCTGTGAGGACCTG 1380
DB 1497 GACAAGGTGACCGCCAGGAGGAGAACTTCTGCCCAAGTACCAAGGCTGTGAGGACCTG 1556
QY 1381 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAAAGTGTGTG 1440
DB 1557 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAAAGTGTGTG 1616
QY 1441 GAGGACGCGGAGGAGCTGAGCTGCATAGTGCAGAGGCGCCCATGCGGCTGGCGCGC 1500
DB 1617 GAGGACGCGGAGGAGCTGAGCTGCATAGTGCAGAGGCGCCCATGCGGCTGGCGCGC 1676
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGCGGCGAGGCGAGCGGCTGCACC 1560
DB 1677 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGCGGCGAGGCGAGCGGCTGCACC 1736
QY 1561 TGTGACAGCGGGAATAAAGCTCAGCTGCGCGGAGCGCGGAAAAAATCTTTCAAGAAG 1620
DB 1737 TGTGACAGCGGGAATAAAGCTCAGCTGCGCGGAGCGCGGAAAAAATCTTTCAAGAAG 1796
QY 1621 AAGTACAGGCGAGCTATGTCGCGAGTCCATCGCTCAGTGGCCATCGAGGTGGAC 1680
DB 1797 AAGTACAGGCGAGCTATGTCGCGAGTCCATCGCTCAGTGGCCATCGAGGTGGAC 1856
QY 1681 GGCAAGGTGTACCAAGTGGGTGATGCCGCCAGCGCCGAAACCTTCAACAGCGG 1740
DB 1857 GGCAAGGTGTACCAAGTGGGTGATGCCGCCAGCGCCGAAACCTTCAACAGCGG 1916
QY 1741 CACTGGCGAGGCGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1800
DB 1917 CACTGGCGAGGCGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1976
QY 1801 GGAGGCTTCCGACTACTCAGCGCGCAACCCCATTTAAAGTGACATCGGTCTACATC 1860
DB 1977 GGAGGCTTCCGACTACTCAGCGCGCAACCCCATTTAAAGTGACATCGGTCTACATC 2036
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACTGGACTGTAAGTCCCTGCAAGGCTCGGAAA 1920
DB 2037 CTAGAGAACGACACAGTCCAGTGTGACTGGACTGTAAGTCCCTGCAAGGCTCGGAAA 2096
QY 1921 GACCACAAGCTGCATCGACCGAGATGAAACCTGCGAGAACAAATTAAGAACCTG 1980
DB 2097 GACCACAAGCTGCATCGACCGAGATGAAACCTGCGAGAACAAATTAAGAACCTG 2156
QY 1981 AGGGAAGTCCGAGTCACTGAAAGAAAGCGGCCAGAGAAATGTGACTGTCACAAAATC 2040
DB 2157 AGGGAAGTCCGAGTCACTGAAAGAAAGCGGCCAGAGAAATGTGACTGTCACAAAATC 2216
QY 2041 AGCTTACCACCCAGCACAAAGCGCGCTCAAGCACAGAGGCTCCAGTCTGCATCTTTC 2100

DB 2217 AGCTTACCACCCAGCACAAAGCGCGCTCAAGCACAGAGGCTCCAGTCTGCATCTTTC 2276
QY 2101 AGGAAGGCGCTGCAAGGAGAGGCAAGGCTGTGCTGTTGGGGAGAGAGCGCAAGAAG 2160
DB 2277 AGGAAGGCGCTGCAAGGAGAGGCAAGGCTGTGCTGTTGGGGAGAGAGCGCAAGAAG 2336
QY 2161 AAACTCCGCAAGCTGTCTCAAGCGCTTCAGAGCAAAACACACAGCTGCAGCATGCCAGGCTC 2220
DB 2337 AAACTCCGCAAGCTGTCTCAAGCGCTTCAGAGCAAAACACACAGCTGCAGCATGCCAGGCTC 2396
QY 2221 ACCTGCTTACCCACGACCAACAGCACTGGGACGCGGCTTCTGGACACTGGGCGCT 2280
DB 2397 ACCTGCTTACCCACGACCAACAGCACTGGGACGCGGCTTCTGGACACTGGGCGCT 2456
QY 2281 TTCTGTCTGCTGACACGCGCCCAACAATAACACAGTACTGTGTCATGAGGACCATCAATGAG 2340
DB 2457 TTCTGTCTGCTGACACGCGCCCAACAATAACACAGTACTGTGTCATGAGGACCATCAATGAG 2516
QY 2341 ACTCAAAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2400
DB 2517 ACTCAAAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2576
QY 2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTCAACAG 2460
DB 2577 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTCAACAG 2636
QY 2461 CTACAGCTACAGCTCATGAGCTGAGAGCTGCAAGGTTTAAAGCAGTGTAAACCCCGG 2520
DB 2637 CTACAGCTACAGCTCATGAGCTGAGAGCTGCAAGGTTTAAAGCAGTGTAAACCCCGG 2696
QY 2521 ACTCGAAACATGAGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGGACGTTT 2580
DB 2697 ACTCGAAACATGAGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGGACGTTT 2756
QY 2581 CAGCGTCGAAAGTGGCAGAAATGAAGAGACCTTTCTTCAAATCACTGGGACAACTGTGG 2640
DB 2757 CAGCGTCGAAAGTGGCAGAAATGAAGAGACCTTTCTTCAAATCACTGGGACAACTGTGG 2816
QY 2641 GAAAGCTGGGAAGGTTAAGAAACACAGAGTGGACCTCCAAACACATAGAGGATCAAC 2700
DB 2817 GAAAGCTGGGAAGGTTAAGAAACACAGAGTGGACCTCCAAACACATAGAGGATCAAC 2876
QY 2701 TGACTGCAAGCGCAATGAAACCATGTGGTGTATTTCCAGCAGACCTGTGTATTTGGCC 2760
DB 2877 TGACTGCAAGCGCAATGAAACCATGTGGTGTATTTCCAGCAGACCTGTGTATTTGGCC 2936
QY 2761 AGGAGGCTGAGAAAGCAAGCACTCTCAGTCAACATGACAGATTTGGAGGATTAAC 2820
DB 2937 AGGAGGCTGAGAAAGCAAGCACTCTCAGTCAACATGACAGATTTGGAGGATTAAC 2996
QY 2821 CAGCAGGAGCAGAGATAACTTCAGAAAGTCCATTTTGGCCCTGCTTTTGGTGGATTA 2880
DB 2997 CAGCAGGAGCAGAGATAACTTCAGAAAGTCCATTTTGGCCCTGCTTTTGGTGGATTA 3056
QY 2881 TACTCTCAGCTGCACAAATGCAATTTTCTGATCAAAAGTCAACCACTTAAACCTTCC 2940
DB 3057 TACTCTCAGCTGCACAAATGCAATTTTCTGATCAAAAGTCAACCACTTAAACCTTCC 3116
QY 2941 CCAGAGCTCAAAAGGAAACCGAGAGCGGAGAGATTTCTTGGAAATTTCT 3000
DB 3117 CCAGAGCTCAAAAGGAAACCGAGAGCGGAGAGATTTCTTGGAAATTTCT 3176
QY 3001 CCCAAGGCGGAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAGCAGTCTCTGTCTTAAA 3060
DB 3177 CCCAAGGCGGAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAGCAGTCTCTGTCTTAAA 3236
QY 3061 TCCTCTTATTTCTTTTGGTTGTCAAAAGAGGAACTTAAGAGCAGGACGAGGCAACGT 3120
DB 3237 TCCTCTTATTTCTTTTGGTTGTCAAAAGAGGAACTTAAGAGCAGGACGAGGCAACGT 3296
QY 3121 GGAGAGGCTGAAACAGTGCAGAGACGTTTGCATTAATGAGTCAAGTGCAGCAAAAGAGATGA 3180

Db 3297 GGAGAGGCTGAAACAGTGCAGAGACGCTTTGACAAATGAGTCACTGACGACCAAAAGAGATGA 3356
Qy 3181 CATTTACCTAGCACTATAAACCCCTGGTTGCTCTGAGAACTGCCTTCATTGTATATAT 3240
Db 3357 CATTTACCTAGCACTATAAACCCCTGGTTGCTCTGAGAACTGCCTTCATTGTATATAT 3416
Qy 3241 GTGACTATTTTACATGTAATCAACATGGAACCTTTTATAGGGGAACTTAATAGAAATCCCAA 3300
Db 3417 GTGACTATTTTACATGTAATCAACATGGAACCTTTTATAGGGGAACTTAATAGAAATCCCAA 3476
Qy 3301 TTTTCAGGAGTGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAA 3352
Db 3477 TTTTCAGGAGTGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAA 3528

RESULT 5
ABN86354
ID ABN86354 standard; cDNA; 4397 BP.
AC AC
XX AC
XX AC
DT 08-OCT-2002 (first entry)
XX Human KIAA1427 polypeptide encoding cDNA.
XX Human; sulphatase; 22437; cytostatic; vulnarary; neuroprotective;
KW gene therapy; KIAA1427; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
PH CDS 286..2943
FT /*tag= a
FT /product= "KIAA1427"
FT /partial

XX WO200252019-A2.

XX 04-JUL-2002.

XX 03-OCT-2001; 2001WO-US30856.

XX 21-DEC-2000; 2000US-257082P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Gluckemann MA, Rudolph-Owen LA;

XX WI; 2002-566677/60.

XX P-PSDB; ABB80922.

XX Identifying agents for modulating (e.g. inhibiting) e.g. tumor
XX establishment, growth or metastases, neuron growth, or wound healing by
XX determining whether a test compound binds with a 22437 polypeptide
XX (human sulfatase) -

XX Disclosure; Fig 3A-S; 143pp; English.

XX The invention relates to identifying a compound useful for modulating at
XX least one phenomenon (e.g. tumor establishment, tumour growth, tumour
XX metastases, epithelial and/or endothelial cell proliferation, neuronal
XX cell growth, wound healing or cerebral injury). The method involves
XX determining whether a test compound binds with a 22437 polypeptide.
XX The identified modulators of 22437 nucleic acid and polypeptide are also
XX useful for treating cancer or wounds (e.g. stroke-related cerebral
XX ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid
XX and polypeptide are useful for diagnosing, preventing or treating a
XX subject having cancer or a cellular proliferation and/or differentiation
XX disorder or at risk of developing cancer or a cellular proliferation
XX and/or differentiation disorder. The present sequence represents a human
XX KIAA1427 polypeptide encoding cDNA, used in alignment studies with the
XX human sulphatase polypeptide, 22437 cDNA.

SQ Sequence 4397 BP; 1192 A; 1173 C; 1121 G; 911 T; 0 other;
Query Match 99.58; Score 3347.2; DB 24; Length 4397;
Best Local Similarity 99.94; Pred. No. 0;
Matches 3349; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GACTCCCGCATCCCAAAAGAACACACAGATCAGCAAAAGAGATGGGCCCCCGGAGC 60
Db GACTCCCGCATCCCAAAAGAACACACAGATCAGCAAAAGAGATGGGCCCCCGGAGC 345
Qy 61 CTCGTCTGTCTGCTGCTGCGCACTGTCTTCTCCTGCTGGTGGAGCTGGCCCTTC 120
Db CTCGTCTGTCTGCTGCTGCGCACTGTCTTCTCCTGCTGGTGGAGCTGGCCCTTC 405
Qy 121 CTGTGCGACACACCGCTGAAAGGCGAGGTTTTCAGAGGACCGCAGGAAATCCCGCCCAAC 180
Db CTGTGCGACACACCGCTGAAAGGCGAGGTTTTCAGAGGACCGCAGGAAATCCCGCCCAAC 465
Qy 181 ATCATCTGTGTGTGACGAGCGACCGAGGATGTGGAGCTGGGTTCCATGCAAGTGTGAAC 240
Db ATCATCTGTGTGTGACGAGCGACCGAGGATGTGGAGCTGGGTTCCATGCAAGTGTGAAC 525
Qy 241 AAGACCGGCGCATCATGAGCGAGCGGCGGCGCACTTCATCAACGCTTCGTGACCAACA 300
Db AAGACCGGCGCATCATGAGCGAGCGGCGGCGCACTTCATCAACGCTTCGTGACCAACA 585
Qy 301 CCCATGTGCTGCCCTCACGCTCCTCATCTCTGCAAGTGTGCAAGTGTGCAAGTGTGCA 360
Db CCCATGTGCTGCCCTCACGCTCCTCATCTCTGCAAGTGTGCAAGTGTGCAAGTGTGCA 645
Qy 361 ACCTACACCAACAAATGAGAACTGCTCGCCCTCTGCGAGGCGACAGCAGAGCGCGC 420
Db ACCTACACCAACAAATGAGAACTGCTCGCCCTCTGCGAGGCGACAGCAGAGCGCGC 705
Qy 421 ACCTTTGCGGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480
Db ACCTTTGCGGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 765
Qy 481 AATGAATACACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db AATGAATACACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825
Qy 541 AACTCCCGCTTTTATAACTACACGCTGTGCGAAACGGGGTGAAGAAAGACAGCGCTCC 600
Db AACTCCCGCTTTTATAACTACACGCTGTGCGAAACGGGGTGAAGAAAGACAGCGCTCC 885
Qy 601 GACTACTCCAAAGATTTACCTCAGACCTCATCAACCAATGACACCGTGTGCTTCTCGC 660
Db GACTACTCCAAAGATTTACCTCAGACCTCATCAACCAATGACACCGTGTGCTTCTCGC 945
Qy 661 ACGTCCAAAGAGATGTACCCGACAGCGCAGTCTCATGCTGCTCATGCTGCTGCTGCTGCT 720
Db ACGTCCAAAGAGATGTACCCGACAGCGCAGTCTCATGCTGCTCATGCTGCTGCTGCTGCT 1005
Qy 721 CACGCGCTGTAGGATTCAGCGCCCAATATTCAGCGCTCTTCCCAAAAGCAATCTCAGAC 780
Db CACGCGCTGTAGGATTCAGCGCCCAATATTCAGCGCTCTTCCCAAAAGCAATCTCAGAC 1065
Qy 781 ATCAGCGCGAGCTACACTACGCGCCCAACCGGACAAACACTGATCATGCTGCTGCTGCT 840
Db ATCAGCGCGAGCTACACTACGCGCCCAACCGGACAAACACTGATCATGCTGCTGCTGCTGCT 1125
Qy 841 GGGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGTCTCCAGCGGAAGCGTTGAG 900
Db GGGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGTCTCCAGCGGAAGCGTTGAG 1185
Qy 901 ACCCTCATGTGGTGGAGCTCCATGAGAGCAATTTACACATGCTGCTGCTGCTGCTGCTGCT 960
Db ACCCTCATGTGGTGGAGCTCCATGAGAGCAATTTACACATGCTGCTGCTGCTGCTGCTGCT 1245
Qy 961 GAGCTGGACAAACAGTATCATGCTATACACCGCCGACCAACCGTGTACCATGCTGCTGCT 1020
Db GAGCTGGACAAACAGTATCATGCTATACACCGCCGACCAACCGTGTACCATGCTGCTGCTGCT 1305

QY 1021 GGCTGTGTAAGGGAATCCATCCATATGAGTTTGACATCAGGGTCCGTTTACGTG 1080
Db 1306 GGCTGTGTAAGGGAATCCATCCATATGAGTTTGACATCAGGGTCCGTTTACGTG 1365
QY 1081 AGGGGCCCAACGCTGGAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTACCTG 1140
Db 1366 AGGGGCCCAACGCTGGAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTTACCTG 1425
QY 1141 GCGCCACACATCTGTGACATTTGAGGCTTGGACATACCTTGGGATATGACGGGAATCC 1200
Db 1426 GCGCCACACATCTGTGACATTTGAGGCTTGGACATACCTTGGGATATGACGGGAATCC 1485
QY 1201 ATCTCAAGCTGTGTGACACGGAGCGGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1260
Db 1486 ATCTCAAGCTGTGTGACACGGAGCGGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1545
QY 1261 AGGGTCTGGGGACTCTCTTTGGTGGAGAGGCAAGCTGTACACAGAGACAAAT 1320
Db 1546 AGGGTCTGGGGACTCTCTTTGGTGGAGAGGCAAGCTGTACACAGAGACAAAT 1605
QY 1321 GACAAAGTGAACCCAGGAGGAGAACTTTCTGCCAAGTACACAGCTGTGAAGACCTG 1380
Db 1606 GACAAAGTGAACCCAGGAGGAGAACTTTCTGCCAAGTACACAGCTGTGAAGACCTG 1665
QY 1381 TGTACAGCTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAASTGGCAGTGTGTG 1440
Db 1666 TGTACAGCTGTGAGTACCAAGCGGCTGTGAGCAGCTGGGACAGAASTGGCAGTGTGTG 1725
QY 1441 GAGGAGCCACGGGGAAGCTGAGCTGCATAGTGAAGGGCCCATCGGCTGGGGCGG 1500
Db 1726 GAGGAGCCACGGGGAAGCTGAGCTGCATAGTGAAGGGCCCATCGGCTGGGGCGG 1785
QY 1501 AGCAGAGCCCTCTCCAACTCTGTGCCAAGTACTACGGGACAGGAGCGGCTGTCCAC 1560
Db 1786 AGCAGAGCCCTCTCCAACTCTGTGCCAAGTACTACGGGACAGGAGCGGCTGTCCAC 1845
QY 1561 TGTGACAGCGGGAAGTACAAGCTGACGCTTGGCGGAGCGCGGGAAGAACTCTTCAAGAG 1620
Db 1846 TGTGACAGCGGGAAGTACAAGCTGACGCTTGGCGGAGCGCGGGAAGAACTCTTCAAGAG 1905
QY 1621 AAGTACAAGSCCAGTATGTCCGAGTGCCTCCATCCGCTCAGTGGCCATCAGGTGGAC 1680
Db 1906 AAGTACAAGSCCAGTATGTCCGAGTGCCTCCATCCGCTCAGTGGCCATCAGGTGGAC 1965
QY 1681 GGAGGGTGTACACGATGAGCCTGGGTGTGCGGCCAGCCCGAACTCTCAAGCGG 1740
Db 1966 GGAGGGTGTACACGATGAGCCTGGGTGTGCGGCCAGCCCGAACTCTCAAGCGG 2025
QY 1741 CACTGGCCAGGGCCGCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAAGTGGCACT 1800
Db 2026 CACTGGCCAGGGCCGCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAAGTGGCACT 2085
QY 1801 GGAGGGCTTCCGACTACTCAGCGGCCAAACCCATTAAGTGAACATCGGTGTACATC 1860
Db 2086 GGAGGGCTTCCGACTACTCAGCGGCCAAACCCATTAAGTGAACATCGGTGTACATC 2145
QY 1861 CTAGAGAACGACACAGTCTCAGTGTGACCTGTATCAAGTCCCTGAGGCTGTGAAA 1920
Db 2146 CTAGAGAACGACACAGTCTCAGTGTGACCTGTATCAAGTCCCTGAGGCTGTGAAA 2205
QY 1921 GACCACAGCTGCACATCGACACGAGATTGAACCTTCAGAAACAAATTAAGAACCTG 1980
Db 2206 GACCACAGCTGCACATCGACACGAGATTGAACCTTCAGAAACAAATTAAGAACCTG 2265
QY 1981 AGGGAAGTCCGAGGTCACTGTGAAGAAAAAGCGCCAGAGAAGATGTGACTGTCAAAAATC 2040
Db 2266 AGGGAAGTCCGAGGTCACTGTGAAGAAAAAGCGCCAGAGAAGATGTGACTGTCAAAAATC 2325
QY 2041 AGCTTACACACCCAGCACAAAGCGCCCTCAAGCACAGAGGCTCCAGTCTGATCTTTTC 2100
Db 2326 AGCTTACACACCCAGCACAAAGCGCCCTCAAGCACAGAGGCTCCAGTCTGATCTTTTC 2385

QY 2101 AGGAAGGCGCTGCAAGAGAAGGACAAAGTGTGCTGTGGGAGCAGAGCGCAAGAAG 2160
Db 2386 AGGAAGGCGCTGCAAGAGAAGGACAAAGTGTGCTGTGGGAGCAGAGCGCAAGAAG 2445
QY 2161 AAATCTCGCAAGCTGTCAAGCGCTTCAGAACAAAGACAGTGCAGCATGCGAGGCTC 2220
Db 2446 AAATCTCGCAAGCTGTCAAGCGCTTCAGAACAAAGACAGTGCAGCATGCGAGGCTC 2505
QY 2221 ACGTGTCTTCAACCAAGCAACAGCACTGGCAGACGCGCTTTCTGGACACTGGGSCCT 2280
Db 2506 ACGTGTCTTCAACCAAGCAACAGCACTGGCAGACGCGCTTTCTGGACACTGGGSCCT 2565
QY 2281 TTCTGTGCTCTGCACAGCGCCAAATAAACAGTACTGTGTGATGAGGACCAATCAATGAG 2340
Db 2566 TTCTGTGCTCTGCACAGCGCCAAATAAACAGTACTGTGTGATGAGGACCAATCAATGAG 2625
QY 2341 ACTCACAAATTTCTCTCTGTGTGAATTTGGAACTGGCTTCTCTAGAGTACTTTGATCTCAAC 2400
Db 2626 ACTCACAAATTTCTCTCTGTGTGAATTTGGAACTGGCTTCTCTAGAGTACTTTGATCTCAAC 2685
QY 2401 ACAGACCCCTACAGAGCTGATGAATGCAATGCAAGCACTGACAGGGATGTCTCAACACAG 2460
Db 2686 ACAGACCCCTACAGAGCTGATGAATGCAATGCAAGCACTGACAGGGATGTCTCAACACAG 2745
QY 2461 CTACACCTTACAGCTCATGAGAGCTGCAAGGCTTCAAGCACTGTAAACCCCGG 2520
Db 2746 CTACACCTTACAGCTCATGAGAGCTGCAAGGCTTCAAGCACTGTAAACCCCGG 2805
QY 2521 ACTCGAAACATGAGACCTGGGACTTAAAGATGAGGAGGATGATGAGCAATACAGGCACTTT 2580
Db 2806 ACTCGAAACATGAGACCTGGGACTTAAAGATGAGGAGGATGATGAGCAATACAGGCACTTT 2865
QY 2581 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
Db 2866 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2925
QY 2641 GAAGGCTGGGAAGGTTAAGAAACAAAGAGGCTGCACTCCAAAGACATAGAGGACATCAC 2700
Db 2926 GAAGGCTGGGAAGGTTAAGAAACAAAGAGGCTGCACTCCAAAGACATAGAGGACATCAC 2985
QY 2701 TGACTGCACAGGCAATGAAGAAACCAATGTGGTGTGATTTCCAGCAGACCTGTGTATTGGCC 2760
Db 2986 TGACTGCACAGGCAATGAAGAAACCAATGTGGTGTGATTTCCAGCAGACCTGTGTATTGGCC 3045
QY 2761 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3046 AGGAGGCTGAGAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3105
QY 2821 CAGCAGGAGCAGAGATAAATTTAGGAAAGTCCATTTTCCCTGCTTTTCTTGGATTA 2880
Db 3106 CAGCAGGAGCAGAGATAAATTTAGGAAAGTCCATTTTCCCTGCTTTTCTTGGATTA 3165
QY 2881 TACTCTACAGCTGCACAAATGCAATTTTGTATCAAAAAAGTCAACCAATACCTCTCCC 2940
Db 3166 TACTCTACAGCTGCACAAATGCAATTTTGTATCAAAAAAGTCAACCAATACCTCTCCC 3225
QY 2941 CCAGAGCTCACAAGAAACCGGAGAGCGGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3226 CCAGAGCTCACAAGAAACCGGAGAGCGGAGAGAGATTTCTTGGAAATTTCT 3285
QY 3001 CCCAAGGCGCAAGAGTCAATTTGAAATTTTAAATCATAGGGGAAAGCAGTCTCTTCTTAAA 3060
Db 3286 CCCAAGGCGCAAGAGTCAATTTGAAATTTTAAATCATAGGGGAAAGCAGTCTCTTCTTAAA 3345
QY 3061 TCCTCTTATTTCTTTGTTGTCAAAAGAGAACTTAAGAGCAGGACAGAGGACAGT 3120
Db 3346 TCCTCTTATTTCTTTGTTGTCAAAAGAGAACTTAAGAGCAGGACAGAGGACAGT 3405
QY 3121 GGAGAGCTGAAAAACAGTGCAGAGACGTTTGAATAATGAGTCACTAGCACAAAAGAGATGA 3180
Db 3406 GGAGAGCTGAAAAACAGTGCAGAGACGTTTGAATAATGAGTCACTAGCACAAAAGAGATGA 3465
QY 3181 CATTTACTAGCACTATAAACCCCTGGTGGCTCTGAAAGAAACCTGCTTATTTATATAT 3240

Db 3466 CATTACCTAGACTATTAACCTCTGCTCCCTGAGAAATGCTTCATTGTATAT 3525
Qy 3241 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATAGAAATCCCAA 3300
Db 3526 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATAGAAATCCCAA 3585
Qy 3301 TTTTTCAGGAGTGTGTGTCAATAAAGCGTCTGTGGCCAGTGTAAAGAAAA 3352
Db 3586 TTTTTCAGGAGTGTGTGTCAATAAAGCGTCTGTGGCCAGTGTAAAGAAAA 3637

RESULT 6
ABS56559
ID ABS56559 standard; cDNA; 4286 BP.

AC ABS56559;
XX 20-JAN-2003 (first entry)
XX Full-length human SULF2 cDNA.

XX Human; SULF2; ss: gene; glucosamine-6-sulphatase; cancer; ischaemia;
KW tumour; angiogenesis; coronary; carotid; arterial occlusive disease;
KW peripheral arterial disease; atherosclerosis; myointimal hyperplasia;
KW thromboangitis obliterans; thrombotic disorder; vasculitis;
KW heart attack; myocardial infarction; vascular death; inflammation;
KW rheumatoid arthritis; asthma; adult respiratory distress syndrome;
KW sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;
KW allograft rejection; lymphoma; thrombosis; sulphatase.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 677..3289
FT CDS /*tag= a
FT /*product= "Human SULF2 protein"
XX W0200259327-A2.

XX 01-AUG-2002.
XX 26-DEC-2001; 2001WO-US49793.
XX 27-DEC-2000; 2000US-258577P.
XX 09-FEB-2001; 2001US-267831P.
XX (REGC) UNIV CALIFORNIA.

XX Rosen SD, Hemmerich S, Tomita M, Palmeri D;
XX WPI; 2002-636541/68.
XX P-PSDB; ABG71836.

XX New sulfatase polypeptides having glucosamine-6-sulfatase activity,
PT useful in screening, discovery and preparation of diagnostic and
PT therapeutic agents for treating cancer, ischemic conditions,
PT inflammation, or thrombosis.

XX Example 4; Fig 10A; 293pp; English.

XX This invention relates to the DNA and protein sequences of a novel
CC polypeptide having glucosamine-6-sulphatase activity. The sulphatases of
CC are useful in screening, discovery and preparation of diagnostic and
CC therapeutic agents for treating cancer, ischaemic conditions,
CC inflammation, or thrombosis. The nucleic acids are useful in preparing
CC the sulfatase polypeptides, identifying the expression of genes in a
CC biological specimen, or generating transgenic non-human animals or
CC site-specific gene modification in cell lines. The host cells are
CC useful in replicating and/or expressing the polynucleotides or nucleic
CC acids. The agents are useful in treating the disorders cited above by
CC reducing tumour growth, inflammation, and thrombosis, or increasing
CC angiogenesis, e.g. by treating coronary, carotid, or arterial occlusive

CC disease, peripheral arterial disease, atherosclerosis, myointimal
CC hyperplasia, thromboangitis obliterans, thrombotic disorders,
CC vasculitis, or preventing ischemic conditions, heart attack (myocardial
CC infarction), or other vascular death. The sulphatases and/or agents are
CC also useful in treating rheumatoid arthritis, asthma, adult respiratory
CC distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple
CC sclerosis, allograft rejection, and spread of lymphomas to cutaneous
CC sites. The present sequence represents the full-length cDNA encoding the
CC mouse SULF2 protein of the invention.

XX
SQ Sequence 4286 BP; 1188 A; 1145 C; 1108 G; 845 T; 0 other;

Query Match 99.5%; Score 3346.8; DB 24; Length 4286;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GACTCCCGCATCCCAAGAGAGCAGCAGATCAGCAAAAAGAGATGGGCCCCCGAGC 60
Db 632 GACTCCCGCATCCCAAGAGAGCAGCAGATCAGCAAAAAGAGATGGGCCCCCGAGC 691
Qy 61 CTGCTGCTGTGCTTGTCTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCTTC 120
Db 692 CTGCTGCTGTGCTTGTCTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCTTC 751
Qy 121 CTGTGCGACACCGCTTGAAGGAGGTTTTCAGAGGACCGGAGGACATCCGCCCAAC 180
Db 752 CTGTGCGACACCGCTTGAAGGAGGTTTTCAGAGGACCGGAGGACATCCGCCCAAC 811
Qy 181 ATCATCTGTGTGTGCTGACGAGCAGCAGGATGTGGAGTGTGGTTCATGCAAGGTGATGAAC 240
Db 812 ATCATCTGTGTGTGCTGACGAGCAGCAGGATGTGGAGTGTGGTTCATGCAAGGTGATGAAC 871
Qy 241 AAGACCCGGCGCATATGAGAGCGGGCGGGCGGCATTTATCAACGGCTTCGTGACCA 300
Db 872 AAGACCCGGCGCATATGAGAGCGGGCGGGCGGCATTTATCAACGGCTTCGTGACCA 931
Qy 301 CCCATGTGTGCGCCCTCAGCGCTCTCATCTCTCATCTCTCAACGGAAGTACGTCCACACCAAC 360
Db 932 CCCATGTGTGCGCCCTCAGCGCTCTCATCTCTCATCTCTCAACGGAAGTACGTCCACACCAAC 991
Qy 361 ACCTACACCAACATGAGAACTGCTCTCGCCCTCTCTGGCAGGACAGCAGAGAGCGCG 420
Db 992 ACCTACACCAACATGAGAACTGCTCTCGCCCTCTCTGGCAGGACAGCAGAGAGCGCG 1051
Qy 421 ACCTTTCGGTGTACCTCAATAGCAGTGTGCGGACAGCTTTCTTCGGGAAGTATCTT 480
Db 1052 ACCTTTCGGTGTACCTCAATAGCAGTGTGCGGACAGCTTTCTTCGGGAAGTATCTT 1111
Qy 481 AATGAATACAAACGGCTCTCTACGTGCCACCGCGGTGGAAGAGTGGGTTCGACTCTCTTAAA 540
Db 1112 AATGAATACAAACGGCTCTCTACGTGCCACCGCGGTGGAAGAGTGGGTTCGACTCTCTTAAA 1171
Qy 541 AACTCCCGCTTTTATACTACAGCTGTGTGGAAAGGGGTGAAGAAAGCAGCGCTCC 600
Db 1172 AACTCCCGCTTTTATACTACAGCTGTGTGGAAAGGGGTGAAGAAAGCAGCGCTCC 1231
Qy 601 GACTACTCCCAAGGATTACCTCACAGACCTCATACCAATGACAGCGGTGAGCTTCTTCGCG 660
Db 1232 GACTACTCCCAAGGATTACCTCACAGACCTCATACCAATGACAGCGGTGAGCTTCTTCGCG 1291
Qy 661 ACGTCCCAAGAGATGTACCCGACAGGCCAGTCTCATGTGTATCAGCCATGAGCGCCCC 720
Db 1292 ACGTCCCAAGAGATGTACCCGACAGGCCAGTCTCATGTGTATCAGCCATGAGCGCCCC 1351
Qy 721 CACGGCCCTGAGGATTACAGCCCAATATTACGCGCTCTTCCCAACAGCATCTCAGCAC 780
Db 1352 CACGGCCCTGAGGATTACAGCCCAATATTACGCGCTCTTCCCAACAGCATCTCAGCAC 1411
Qy 781 ATCAGCGCGAGCTCAACTAGCGGCCCAACCCGGAACAAACATGGATCATGGCGGTACAG 840
Db 1412 ATCAGCGCGAGCTCAACTAGCGGCCCAACCCGGAACAAACATGGATCATGGCGGTACAG 1471
Qy 841 GGGCCCATGAAGCCCATCCATCGAATTTCACCAACATGCTCCAGCGGAAGCGGCTTCAG 900

1472 GGGCCCATGAAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAAGCGCTTGCGAG 1531
901 ACCCTCATGTCGGTGGAGCACTCCATGGAGAGGATTTACACATGCTGGTTGAGACGGGC 960
1532 ACCCTCATGTCGGTGGAGCACTCCATGGAGAGGATTTACACATGCTGGTTGAGACGGGC 1591
961 GAGCTGGACAAACACGTPACATCGTATACACCGCGGACCGGTTACACATCGGCGCAGTTT 1020
1592 GAGCTGGACAAACACGTPACATCGTATACACCGCGGACCGGTTACACATCGGCGCAGTTT 1651
1021 GGCCTGGTGAAGGGAATCCATGCGCATATGATGTTGACATCAGGGTCCCGTTCTACGTG 1080
1652 GGCCTGGTGAAGGGAATCCATGCGCATATGATGTTGACATCAGGGTCCCGTTCTACGTG 1711
1081 AGGGGCCCCAAGCTGGGAAGCCGGCTGTCTGAATCCCGACATCGTCTCAACATTTGACCTG 1140
1712 AGGGGCCCCAAGCTGGGAAGCCGGCTGTCTGAATCCCGACATCGTCTCAACATTTGACCTG 1771
1141 GCGCCACCATCTGTGACATTTGACGCTTGACATACCTGCGGATATGAGCGGGAATTC 1200
1772 GCGCCACCATCTGTGACATTTGACGCTTGACATACCTGCGGATATGAGCGGGAATTC 1831
1201 ATCTCAAGCTCTGTGACACGAGCGCGGCTGAATCGGTTTCACTTCAAAAAGAGATG 1260
1832 ATCTCAAGCTCTGTGACACGAGCGCGGCTGAATCGGTTTCACTTCAAAAAGAGATG 1891
1261 AGGGTCTGGCGGACTCTCTTCTGTTGGAGAGAGCAAGCTGTCTACACAGAGAGACAAT 1320
1892 AGGGTCTGGCGGACTCTCTTCTGTTGGAGAGAGCAAGCTGTCTACACAGAGAGACAAT 1951
1321 GACAAAGTGGAGCGCCAGAGGAGAACTTCTGCGCCAAAGTACCGAGCGTGTCAAGGACCTG 1380
1952 GACAAAGTGGAGCGCCAGAGGAGAACTTCTGCGCCAAAGTACCGAGCGTGTGAAGGACCTG 2011
1381 TGTACGCTGTGAGTACACAGCGCGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440
2012 TGTACGCTGTGAGTACACAGCGCGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2071
1441 GAGGAGCGCACGGGAGAGCTGAAGCTGATAGTGCRAAGGCGCCCATGCGGCTGGCGGC 1500
2072 GAGGAGCGCACGGGAGAGCTGAAGCTGATAGTGCRAAGGCGCCCATGCGGCTGGCGGC 2131
1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGCGAGGCGAGGCGCTGCACC 1560
2132 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGCGAGGCGAGGCGCTGCACC 2191
1561 TGTGACAGCGGGAATAAGCTCAGCTGCGCGGACCGCGGACCGCGGAACTCTTCAAGAAG 1620
2192 TGTGACAGCGGGAATAAGCTCAGCTGCGCGGACCGCGGACCGCGGAACTCTTCAAGAAG 2251
1621 AAGTACAGGCGCAGTATGTCGCGAGTGCCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 1680
2252 AAGTACAGGCGCAGTATGTCGCGAGTGCCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2311
1681 GGCAGGGTGTACCAAGCTAGGCTGGGTGATGCGCCAGCGCCCGAAACCTCAACAAGCGG 1740
2312 GGCAGGGTGTACCAAGCTAGGCTGGGTGATGCGCCAGCGCCCGAAACCTCAACAAGCGG 2371
1741 CACTGGCCAGGGCCCTGAGGACCAAGATGACAGATGATGGGACTTCAGTGGCACT 1800
2372 CACTGGCCAGGGCCCTGAGGACCAAGATGACAGATGATGGGACTTCAGTGGCACT 2431
1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAAGTGACATCGGTGTACATC 1860
2432 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAAGTGACATCGGTGTACATC 2491
1861 CTAGAGACGACACAGTCCAGTGTGACTGACCTGTGACCTGTGACCTGAGGCTGGAAA 1920
2492 CTAGAGACGACACAGTCCAGTGTGACTGACCTGTGACCTGTGACCTGAGGCTGGAAA 2551
1921 GACCAAGAGCTGCATCGACCAAGAGATTTGAACCCCTGCGAGAACAAATTAAGAACCTG 1980

2552 GACCACAAGCTGCACATCGACCACGAGATTGAAACCCCTGCAGAACAAATTAAGAACCTG 2611
1981 AGGAAAGTCCGAGGTCACTGTAAGAAAACCGCCAGAGAAATGCTGCTGCAAAAATC 2040
2612 AGGAAAGTCCGAGGTCACTGTAAGAAAACCGCCAGAGAAATGCTGCTGCAAAAATC 2671
2041 AGCTACACACACCCAGCACAAAGCCGCCCTCAAGCAAGAGGCTTCCAGTCTGATCCCTTTC 2100
2672 AGCTACACACACCCAGCACAAAGCCGCCCTCAAGCAAGAGGCTTCCAGTCTGATCCCTTTC 2731
2101 AGGAAAGTCCGAGGTCACTGTAAGAAAACCGCCAGAGAAATGCTGCTGCAAAAATC 2160
2732 AGGAAAGTCCGAGGTCACTGTAAGAAAACCGCCAGAGAAATGCTGCTGCAAAAATC 2791
2161 AAAGTCCGAGGCTGCTCAAGCGCCCTGCAAGCAACAGCAACGTCAGCATGCGAGGCTTC 2220
2792 AAAGTCCGAGGCTGCTCAAGCGCCCTGCAAGCAACAGCAACGTCAGCATGCGAGGCTTC 2851
2221 AGTGTCTTCCACCAAGCAACAGCAACGTCAGCATGCGAGGCTTCCTGAGACATCGGGGCT 2280
2852 AGTGTCTTCCACCAAGCAACAGCAACGTCAGCATGCGAGGCTTCCTGAGACATCGGGGCT 2911
2281 TTCTGTGCTGACACAGCGCCCAACATTAACAGTACTGTCATGAGGACCATCAATGAG 2340
2912 TTCTGTGCTGACACAGCGCCCAACATTAACAGTACTGTCATGAGGACCATCAATGAG 2971
2341 ACTCAAAATTTCTTCTGTGTGAATTTGCAATGCGCTTCTTAGAGTACTTGTGATCTCAAC 2400
2972 ACTCAAAATTTCTTCTGTGTGAATTTGCAATGCGCTTCTTAGAGTACTTGTGATCTCAAC 3031
2401 ACAGACCCCTACAGCTGATGATGAGTGAACACACTGAGGAGTATGCTTCAACACAG 2460
3032 ACAGACCCCTACAGCTGATGATGAGTGAACACACTGAGGAGTATGCTTCAACACAG 3091
2461 CTACAGCTCAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGTAAACCCCGG 2520
3092 CTACAGCTCAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGTAAACCCCGG 3151
2521 ACTCGAAAATGAGCCTGGGACTTAAAGTGGAGGAGTATGAGCAATACAGGCGATTT 2580
3152 ACTCGAAAATGAGCCTGGGACTTAAAGTGGAGGAGTATGAGCAATACAGGCGATTT 3211
2581 CAGCGTCGAAAAGTGGGCGCAAAATGAAGAGACCTTCTTCAAAATCACTGGGACCACTGG 2640
3212 CAGCGTCGAAAAGTGGGCGCAAAATGAAGAGACCTTCTTCAAAATCACTGGGACCACTGG 3271
2641 GAAGGCTGGGAGGTTAAGAAAACAGAGGTGGACCTTCAAAAACATAGAGGCGATCAC 2700
3272 GAAGGCTGGGAGGTTAAGAAAACAGAGGTGGACCTTCAAAAACATAGAGGCGATCAC 3331
2701 TGACTGCACAGCAATGAAGAAAACCATGTTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
3332 TGACTGCACAGCAATGAAGAAAACCATGTTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3391
2761 AGGAGGCTTGAGAAAACAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
3392 AGGAGGCTTGAGAAAACAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3451
2821 CAGCAGGACAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880
3452 CAGCAGGACAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3511
2881 TACTCACCAGCTGACAAAATGCTATTTTTCGTATCAAAAAGTCAACCAATCAACCTCC 2940
3512 TACTCACCAGCTGACAAAATGCTATTTTTCGTATCAAAAAGTCAACCAATCAACCTCC 3571
2941 CAGAGAGCTCAAGAGGAAAACGAGAGAGCGAGGAGATTTCTTGGAAAATTTCT 3000
3572 CAGAGAGCTCAAGAGGAAAACGAGAGAGCGAGGAGATTTCTTGGAAAATTTCT 3631
3001 CCAAGGGGCAAGTCAATGGAAATTTTAAATCATAGGGGAAAACGAGTCTGTGTTAAA 3060
3632 CCAAGGGGCAAGTCAATGGAAATTTTAAATCATAGGGGAAAACGAGTCTGTGTTAAA 3691

CC The invention provides human drug metabolizing enzymes (DME) and polynucleotides encoding the DMEs. The DME can be expressed by standard recombinant methodology. DMEs and their agonists and antagonists are useful for the diagnosis, treatment, and prevention of autoimmune/inflammatory, cell proliferative, developmental, endocrine such as aneurysm, eye, metabolic, and gastrointestinal disorders, including liver disorders and infection. The present sequence represents a human DME encoding cDNA.

XX
SQ Sequence 3348 BP; 927 A; 933 C; 869 G; 619 T; 0 other;

Query Match 99.58; Score 3344.8; DB 22; Length 3348;
Best Local Similarity 99.94; Pred. No. 0;
Matches 3346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 12 CCCAAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGCCTCGTGTGTG 71
Db 1 CCCAAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGCCTCGTGTGTG 60
QY 72 CTTGCTGTCGCAACTGTGTTCTCCCTGCTGGGTGAAGCTCGGCTTCTGTGCGACCA 131
Db 61 CTTGCTGTCGCAACTGTGTTCTCCCTGCTGGGTGAAGCTCGGCTTCTGTGCGACCA 120
QY 132 CCGCTCAAAAGGAGGTTTCAGAGGACCGCAGAACATCCGCCCAACATCATCTGTGT 191
Db 121 CCGCTCAAAAGGAGGTTTCAGAGGACCGCAGAACATCCGCCCAACATCATCTGTGT 180
QY 192 GCTGACGGACGACGAGATGTGAGCTGGTTCATGAGGTGATGAAACAGACCCGGCG 251
Db 181 GCTGACGGACGACGAGATGTGAGCTGGTTCATGAGGTGATGAAACAGACCCGGCG 240
QY 252 CATCATGAGCAGGCGGGGCGACATTCATCAAGCCTTCGTGACCAACACCCATGTCTG 311
Db 241 CATCATGAGCAGGCGGGGCGACATTCATCAAGCCTTCGTGACCAACACCCATGTCTG 300
QY 312 CCCCTACGCTCTCCATCTCTGCAAGTACGTCCACCAACCAACACCTACACAA 371
Db 301 CCCCTACGCTCTCCATCTCTGCAAGTACGTCCACCAACCAACACCTACACAA 360
QY 372 CAATGAGAACTGCTCTCCGCTCTGCGAGGACAGCAGGAGCGGACCTTTCGCT 431
Db 361 CAATGAGAACTGCTCTCCGCTCTGCGAGGACAGCAGGAGCGGACCTTTCGCT 420
QY 432 GTACCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTTAATGAATCAA 491
Db 421 GTACCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTTAATGAATCAA 480
QY 492 CGGCTCTTAAGTCCACCGGCTGGAAGAGTGGTGGAGTCTCTTAATAAATCCCGCTT 551
Db 481 CGGCTCTTAAGTCCACCGGCTGGAAGAGTGGTGGAGTCTCTTAATAAATCCCGCTT 540
QY 552 TTATAACTACACGCTGTGCGAAGCGGCTGAAGAAAGCAGCGCTCCGACTACTCCAA 611
Db 541 TTATAACTACACGCTGTGCGAAGCGGCTGAAGAAAGCAGCGCTCCGACTACTCCAA 600
QY 612 GGATTAACCTCAACAGCTCTATCAACAAATGACAGCTGAGTCTTTCGCGAGCTCAAGAA 671
Db 601 GGATTAACCTCAACAGCTCTATCAACAAATGACAGCTGAGTCTTTCGCGAGCTCAAGAA 660
QY 672 GATGTACCGCAGACGCGCTCTCATGCTCATCGCCATGCGACCCCGCCCGCTCGA 731
Db 661 GATGTACCGCAGACGCGCTCTCATGCTCATCGCCATGCGACCCCGCCCGCTCGA 720
QY 732 GGATTCAGCCCCCACAATATTCACGCTCTTCCCAAGCAGTCTCAGACATCAGCGGAG 791
Db 721 GGATTCAGCCCCCACAATATTCACGCTCTTCCCAAGCAGTCTCAGACATCAGCGGAG 780
QY 792 CTACAACTACCGCGCCCAACCGGCAACACTGATCATGGCTTACACGGGGCCCATGAA 851
Db 781 CTACAACTACCGCGCCCAACCGGCAACACTGATCATGGCTTACACGGGGCCCATGAA 840
QY 852 GCCCATCCACATGGAATTCACCAACATGCTCTCAGCGGAAGCGCTTGAGACCCCTCATGTC 911

AAH76201
ID AAH76201 standard; cDNA; 3348 BP.
AC AAH76201;
XX 29-OCT-2001 (first entry)
XX Human drug metabolizing enzyme encoding cDNA (ID No. 1559210CB1).
XX Drug metabolizing enzyme; DME; immunosuppressive; cytostatic; ophthalmic;
KW hepatotropic; antiallergic; antiaesthetic; antibacterial; antiviral;
KW antisense therapy; gene therapy; human; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 35..2647
XX /*tag= a
XX WO200159127-A2.
XX 16-AUG-2001.
XX 08-FEB-2001; 2001WO-US04423.
XX 11-FEB-2000; 2000US-0181856.
XX 17-FEB-2000; 2000US-0183684.
XX 25-FEB-2000; 2000US-0185141.
XX 03-MAR-2000; 2000US-0186818.
XX 09-MAR-2000; 2000US-0188345.
XX 17-MAR-2000; 2000US-0189997.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Lal P;
PI Gandhi AR, Ring HZ, Shih LL, Yang J, Policky JL;
XX WPI; 2001-514673/56.
XX P-PSDB; AAB85774.
XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the
PI diagnosis, treatment, and prevention of autoimmune/inflammatory, cell
XX proliferative, developmental and endocrine disorders -
PS Claim 5; Page 144-145; 150pp; English.
XX

Db 841 GCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTGAGACACCTCATGTC 900
Qy 912 GGTGAGAGACTCCATGGAGACGATTTTAAACATGCTGTTGAGACGCGGAGCTGGACAA 971
Db 901 GGTGAGACACTCCATGGAGACGATTTTAAACATGCTGTTGAGACGCGGAGCTGGACAA 960
Qy 972 CACGTACATCGTATACACCGCGGACCAACGGTTACCAATCGCCAGTTTGGCCTGGTGA 1031
Db 961 CACGTACATCGTATACACCGCGGACCAACGGTTACCAATCGCCAGTTTGGCCTGGTGA 1020
Qy 1032 AGGGAATTCATGCCATATGATTTGACATCAGAGTCCCGTTTCTACGTGAGGGGCCCAA 1091
Db 1021 AGGGAATTCATGCCATATGATTTGACATCAGAGTCCCGTTTCTACGTGAGGGGCCCAA 1080
Qy 1092 CGTGAAGCCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCGCCACCAT 1151
Db 1081 CGTGAAGCCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCGCCACCAT 1140
Qy 1152 CTTGAGACATTTGAGCCCTGGACATACCTGGCGATATGGACGGGAATCCATCTCAAGCT 1211
Db 1141 CTTGAGACATTTGAGCCCTGGACATACCTGGCGATATGGACGGGAATCCATCTCAAGCT 1200
Qy 1212 GCTGACACGAGCGCGCGGTGAATCGGTTTCACTTGAAAAGAGATGAGGGTCTGGCG 1271
Db 1201 GCTGACACGAGCGCGCGGTGAATCGGTTTCACTTGAAAAGAGATGAGGGTCTGGCG 1260
Qy 1272 GGAATCTCTTGTGTGAGAGAGGCAAGCTGCTACAAAGAGACAAATGACAAAGTGGA 1331
Db 1261 GGAATCTCTTGTGTGAGAGAGGCAAGCTGCTACAAAGAGACAAATGACAAAGTGGA 1320
Qy 1332 CGCCAGAGAGAGAACTTTCTGCCCAAGTACACAGCGTGTGAAGACCTGTGTAGCGTGC 1391
Db 1321 CGCCAGAGAGAGAACTTTCTGCCCAAGTACACAGCGTGTGAAGACCTGTGTAGCGTGC 1380
Qy 1392 TGAGTACACAGCGCGTGTGAGAGCTGGGACAGAGTGGCAGTGTGTGGAGGACGCCAC 1451
Db 1381 TGAGTACACAGCGCGTGTGAGAGCTGGGACAGAGTGGCAGTGTGTGGAGGACGCCAC 1440
Qy 1452 GGGGAAGCTGAAGCTGCATAGTGCAGGGGCCCATGCGGCTGGCGGACAGAGCCCT 1511
Db 1441 GGGGAAGCTGAAGCTGCATAGTGCAGGGGCCCATGCGGCTGGCGGACAGAGCCCT 1500
Qy 1512 CTCCAACCTCGTCCCAAGTACTACGGGACGGGACGAGGCTTCACCTGTGACAGCGG 1571
Db 1501 CTCCAACCTCGTCCCAAGTACTACGGGACGGGACGAGGCTTCACCTGTGACAGCGG 1560
Qy 1572 GGAATCAAGCTCAGCTGGCGGACGCGGAAACCTCTTCAAGAGAAAGTACAAGGC 1631
Db 1561 GGAATCAAGCTCAGCTGGCGGACGCGGAAACCTCTTCAAGAGAAAGTACAAGGC 1620
Qy 1632 CAGCTATGTCCGAGTCCGCTCCATCGGCTCAGTGGCCATCGAGTGGACGCGAGGGTGTA 1691
Db 1621 CAGCTATGTCCGAGTCCGCTCCATCGGCTCAGTGGCCATCGAGTGGACGCGAGGGTGTA 1680
Qy 1692 CCACTAGGCTGGGTGATGCGGCCACGCCCGAAACCTCAACAGCGGCACTGGCCAGG 1751
Db 1681 CCACTAGGCTGGGTGATGCGGCCACGCCCGAAACCTCAACAGCGGCACTGGCCAGG 1740
Qy 1752 GGGCCCTCAGGACCAAGATGACAGGATGGTGGGACTTTCAGTGGCACTGGAGGCCCTTC 1811
Db 1741 GGGCCCTCAGGACCAAGATGACAGGATGGTGGGACTTTCAGTGGCACTGGAGGCCCTTC 1800
Qy 1812 CGACTACTCAGCGCCCAACCCCATTAAGTGAACATCGTGTGCTACATCTCTAGAGACGA 1871
Db 1801 CGACTACTCAGCGCCCAACCCCATTAAGTGAACATCGTGTGCTACATCTCTAGAGACGA 1860
Qy 1872 CACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGCCCTGGAAAGACCAAGCT 1931
Db 1861 CACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGCCCTGGAAAGACCAAGCT 1920
Qy 1932 GCACATCCAGCAGAGATTGAACCCCTGCAGAAACAAATTAAGAACCTTGAGGGAGTCCG 1991
Db 1921 GCACATCCAGCAGAGATTGAACCCCTGCAGAAACAAATTAAGAACCTTGAGGGAGTCCG 1980

Qy 1992 AGTTCACCTGGAAGAAAAGCGGCCAGAGAAGTGTGACTGTACAAAAATCAGCTACACAC 2051
Db 1981 AGTTCACCTGGAAGAAAAGCGGCCAGAGAAGTGTGACTGTACAAAAATCAGCTACACAC 2040
Qy 2052 CCAGCACAAGAGCCGCTTCAAGCAGAGGCTCCAGTCTGCAATCTGCAATCTTTTCAAGAAAGGCT 2111
Db 2041 CCAGCACAAGAGCCGCTTCAAGCAGAGGCTCCAGTCTGCAATCTGCAATCTTTTCAAGAAAGGCT 2100
Qy 2112 GCAAGAGAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGCAAGAGAAATCCGCAA 2171
Db 2101 GCAAGAGAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGCAAGAGAAATCCGCAA 2160
Qy 2172 GTGTCTCAAGCCCTTCCAGAAACAACGACACGTCGACGATGCCAGGCTTCACTGTCTTCA 2231
Db 2161 GTGTCTCAAGCCCTTCCAGAAACAACGACACGTCGACGATGCCAGGCTTCACTGTCTTCA 2220
Qy 2232 CCAGCACAAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGGCTTCTTGGACCTGGGGCTTCTGTGGCTG 2291
Db 2221 CCAGCACAAGAGCAAGGTGTGGCTGTTCGGGAGCAGAGCGGCTTCTTGGACCTGGGGCTTCTGTGGCTG 2280
Qy 2292 CACAGCGCCAAACAATAACACGTAAGTGTGATGAGGACCAATCAATGAGACTCACAAATTT 2351
Db 2281 CACAGCGCCAAACAATAACACGTAAGTGTGATGAGGACCAATCAATGAGACTCACAAATTT 2340
Qy 2352 CTTCTTCTGTGAATTTGMAAATGCGTCTTAGAGTACTTTGATCTCAACAGACCCCTA 2411
Db 2341 CTTCTTCTGTGAATTTGMAAATGCGTCTTAGAGTACTTTGATCTCAACAGACCCCTA 2400
Qy 2412 CAGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTTCAACAGCTACAGGTACA 2471
Db 2401 CAGCTGATGAATGCAAGTGAACACACTGGACAGGAGTGTCTTCAACAGCTACAGGTACA 2460
Qy 2472 GTCATGAGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGGACCTCGAAAAAT 2531
Db 2461 GTCATGAGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGGACCTCGAAAAAT 2520
Qy 2532 GGACCTGGAGCTTAAAGATGGAGGAGCTATGAGCAATACAGGCAAGTTTACGCTCGAAA 2591
Db 2521 GGACCTGGAGCTTAAAGATGGAGGAGCTATGAGCAATACAGGCAAGTTTACGCTCGAAA 2580
Qy 2592 GTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGGGAAGGCTGGGA 2651
Db 2581 GTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGGGAAGGCTGGGA 2640
Qy 2652 AGGTTAAGAAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACTGACTGCACAG 2711
Db 2641 AGGTTAAGAAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACTGACTGCACAG 2700
Qy 2712 GCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGGA 2771
Db 2701 GCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGGA 2760
Qy 2772 GAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATCTGGAGGATTAACAGCAGAGCA 2831
Db 2761 GAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATCTGGAGGATTAACAGCAGAGCA 2820
Qy 2832 GAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTTATACCTCACCAG 2891
Db 2821 GAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTTATACCTCACCAG 2880
Qy 2892 CTGCAAAAATGCAATTTTTCGATCAAAAAGTCAACCACTAACCCCTCCCAAGAGCTCA 2951
Db 2881 CTGCAAAAATGCAATTTTTCGATCAAAAAGTCAACCACTAACCCCTCCCAAGAGCTCA 2940
Qy 2952 CAAAGGAAAACGAGAGCGGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGGA 3011
Db 2941 CAAAGGAAAACGAGAGCGGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGGA 3000
Qy 3012 AAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAAATCTCTTATTC 3071
Db 3001 AAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAAATCTCTTATTC 3060

QY 3072 TTTTGGTTTGTCTCAAGAAGGAAGAACTAAGAAGCAGGACAGGCAACGCTGGAGAGGCTGA 3131
Db 3061 TTTTGGTTTGTCTCAAGAAGGAAGAACTAAGAAGCAGGACAGGCAACGCTGGAGAGGCTGA 3120
QY 3132 AAACAGTGCAGAGACGTTTGACAAATGAGTCAGTAGCACAAGAGATGACATTTACCTAG 3191
Db 3121 AAACAGTGCAGAGACGTTTGACAAATGAGTCAGTAGCACAAGAGATGACATTTACCTAG 3180
QY 3192 CACTATAAACCCCTGGTTCCTCTGAAGAACTGCTTCATTGTATATATATGACTATTTA 3251
Db 3181 CACTATAAACCCCTGGTTCCTCTGAAGAACTGCTTCATTGTATATATATGACTATTTA 3240
QY 3252 CATGTAATCAATGGAACCTTTTAGGGGAACCTATAAGAAATCCCAATTTTCAGGAGT 3311
Db 3241 CATGTAATCAATGGAACCTTTTAGGGGAACCTATAAGAAATCCCAATTTTCAGGAGT 3300
QY 3312 GGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAGAAAGAAAAAAA 3359
Db 3301 GGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAGAAAGAAAAAAA 3348

RESULT 8

AAZ64969

ID AAZ64969 standard; cDNA; 3906 BP.

XX AC AAZ64969;

XX XX

DT 05-APR-2000 (first entry)

XX

DE Membrane-bound protein PRO1120 encoding cDNA.

XX

KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;

KW

pharmaceutical; receptor immunoadhesin; gene mapping; ss.

XX XX

OS Homo sapiens.

XX

PN WO9963088-A2.

XX

PD 09-DEC-1999.

XX

PF 02-JUN-1999; 99WO-US12252.

XX

XX 02-JUN-1998; 98US-0087607.

PR

PR 02-JUN-1998; 98US-0087609.

PR

PR 03-JUN-1998; 98US-0087759.

PR

PR 04-JUN-1998; 98US-0087827.

PR

PR 04-JUN-1998; 98US-0088021.

PR

PR 04-JUN-1998; 98US-0088025.

PR

PR 04-JUN-1998; 98US-0088028.

PR

PR 04-JUN-1998; 98US-0088030.

PR

PR 04-JUN-1998; 98US-0088033.

PR

PR 04-JUN-1998; 98US-0088326.

PR

PR 05-JUN-1998; 98US-0088167.

PR

PR 05-JUN-1998; 98US-0088202.

PR

PR 05-JUN-1998; 98US-0088212.

PR

PR 05-JUN-1998; 98US-0088217.

PR

PR 09-JUN-1998; 98US-0088655.

PR

PR 10-JUN-1998; 98US-0088722.

PR

PR 10-JUN-1998; 98US-0088730.

PR

PR 10-JUN-1998; 98US-0088734.

PR

PR 10-JUN-1998; 98US-0088738.

PR

PR 10-JUN-1998; 98US-0088740.

PR

PR 10-JUN-1998; 98US-0088741.

PR

PR 10-JUN-1998; 98US-0088742.

PR

PR 10-JUN-1998; 98US-0088810.

PR

PR 10-JUN-1998; 98US-0088811.

PR

PR 10-JUN-1998; 98US-0088824.

PR

PR 10-JUN-1998; 98US-0088825.

PR

PR 10-JUN-1998; 98US-0088826.

PR

PR 11-JUN-1998; 98US-0088858.

PR

PR 11-JUN-1998; 98US-0088861.

PR

PR 11-JUN-1998; 98US-0088863.

PR

PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 25-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094551.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.

```
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-010634.
PR 12-JAN-1999; 99US-0115565.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
XX Wood WI, Yuan J;
XX
XX WPI; 2000-072883/06.
XX P-PSDB; AAY66648.
XX
XX Membrane-bound proteins and related nucleotide sequences -
XX
XX Claim 2; Fig 46 A-B; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
XX polynucleotides encoding them. The PRO sequences of the invention were
XX identified based on extracellular domain homology screening. The PRO
XX sequences have homology with proteins including LDL receptors, TIE
XX ligands and various enzymes. The membrane-bound proteins and receptor
XX molecules are useful as pharmaceutical and diagnostic agents. Receptor
XX immunoadhesins, for instance, can be used as therapeutic agents to block
XX receptor-ligand interactions. The membrane-bound proteins can also be
XX employed for screening of potential peptide or small molecule inhibitors
XX of the relevant receptor/ligand interaction. The PRO encoding sequences
XX are useful as hybridization probes, in chromosome and gene mapping and in
XX the generation of antisense RNA and DNA. PRO nucleic acid sequences
XX will also be useful for the preparation of PRO polypeptides, especially
XX by recombinant techniques.
XX
XX Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;
XX
XX Query Match 98.8%; Score 3323.4; DB 21; Length 3906;
XX Best Local Similarity 99.7%; Pred. No. 0;
XX Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
XX
XX QY 1 GACTCCCGCATCCCAAAAGAACACAGATCAGCAAAAAGAGATGGCCCCCGGAGC 60
XX
XX Db 563 GACTCCCGCATCCCAAAAGAACACAGATCAGCAAAAAGAGATGGCCCCCGGAGC 622
XX
XX QY 61 CTGCTGCTGTGCTGTGTCGCGCACTGTCTCCCTGCTGGTGGAGCTCGGCCTTC 120
XX
XX Db 623 CTGCTGCTGTGCTGTGTCGCGCACTGTCTCCCTGCTGGTGGAGCTCGGCCTTC 682
```

1201 ATCTCAAGCTGTGACACGAGCGCGGTGAATCGGTTTCACTTGAAGAAGATG 1260
1763 ATCTCAAGCTGTGACACGAGCGCGGTGAATCGGTTTCACTTGAAGAAGATG 1822
1261 AGGGTCTGGCGGACTCCTTCTTGGTGGAGAGGCAAGCTCTACACAAGAGACAAAT 1320
1823 AGGGTCTGGCGGACTCCTTCTTGGTGGAGAGGCAAGCTCTACACAAGAGACAAAT 1882
1321 GACAAAGTGGACGCCACGAGAGGAACTTCTGCCCCAAGTACACGGTGTGAAGACCTG 1380
1883 GACAAAGTGGACGCCACGAGAGGAACTTCTGCCCCAAGTACACGGTGTGAAGACCTG 1942
1381 TGTCAAGCTGTGAGTACACAGCGCGGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440
1943 TGTCAAGCTGTGAGTACACAGCGCGGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2002
1441 GAGGAGCCACGCGGGAAGCTGAAGCTGCATTAAGTGAAGGGGCCCTGCGGCTGGCGGC 1500
2003 GAGGAGCCACGCGGGAAGCTGAAGCTGCATTAAGTGAAGGGGCCCTGCGGCTGGCGGC 2062
1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGCAGGCGAGCGCTGCACC 1560
2063 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGCAGGCGAGCGCTGCACC 2122
1561 TGTGACAGCGGGACTACAAAGCTGAGCTGTGCGCGAGCGCGGAAAACTCTTCAAGAG 1620
2123 TGTGACAGCGGGACTACAAAGCTGAGCTGTGCGCGAGCGCGGAAAACTCTTCAAGAG 2182
1621 AAGTACAGGCCAGCTATGTTCGCGAGTTCGCTCCATCGCTCAAGTGCATCGAGGTGGAC 1680
2183 AAGTACAGGCCAGCTATGTTCGCGAGTTCGCTCCATCGCTCAAGTGCATCGAGGTGGAC 2242
1681 GGCAGGCTGTACACGTAGGCTTGGGTGATGCGCGCCCGCGAACCCTCCAAAGCGG 1740
2243 GGCAGGCTGTACACGTAGGCTTGGGTGATGCGCGCCCGCGAACCCTCCAAAGCGG 2302
1741 CACTGGCCAGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAGTGGCACT 1800
2303 CACTGGCCAGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACTTCAGTGGCACT 2362
1801 GGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGAACATCGGTGTACATC 1860
2363 GGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGAACATCGGTGTACATC 2422
1861 CTAGAGAACACACAGTTCAGTGTGACCTGACCTGTACAGTCCCTCAGGGCTGGAAA 1920
2423 CTAGAGAACACACAGTTCAGTGTGACCTGACCTGTACAGTCCCTCAGGGCTGGAAA 2482
1921 GACCACAGCTGCACATCGACACGAGATTGAAACCTCGCAGAACAAATTAAGAACCTG 1980
2483 GACCACAGCTGCACATCGACACGAGATTGAAACCTCGCAGAACAAATTAAGAACCTG 2542
1981 AGGGAATCGGAGTGTACCTGAGAAAGAACCGGCGCAGAGAAATGTGACTGTCAAAAATC 2040
2543 AGGGAATCGGAGTGTACCTGAGAAAGAACCGGCGCAGAGAAATGTGACTGTCAAAAATC 2602
2041 AGCTTACCACACCCAGCACAAGCGCGCTCAAGACAGAGGCTCCAGTCTGCACTCTTTC 2100
2603 AGCTTACCACACCCAGCACAAGCGCGCTCAAGACAGAGGCTCCAGTCTGCACTCTTTC 2662
2101 AGGAAGCGCTTGCAGAGAGGACCAAGGTGTGCTGTGCGGAGCAGAGCGCAAGAG 2160
2663 AGGAAGCGCTTGCAGAGAGGACCAAGGTGTGCTGTGCGGAGCAGAGCGCAAGAG 2722
2161 AACTCTCGAGCTGTCTCAGCGCTTCAGACCAAGACAGCTGCAGCATGCCAGGCTC 2220
2723 AACTCTCGAGCTGTCTCAGCGCTTCAGACCAAGACAGCTGCAGCATGCCAGGCTC 2782
2221 ACGTGCTTCCACACGACCAACAGCACTGGCAGACGGCGCTTCTGGAACCTGGGCGCT 2280
2783 ACGTGTCTTCCACACGACCAACAGCACTGGCAGACGGCGCTTCTGGAACCTGGGCGCT 2842
2281 TTCTGTGCTGCACACGCGCCCAACATAACAGTACTGTGTGATGAGGACCATCAATGAG 2340

RESULT 9

2843 TTCTGTGCTGCACACGCGCCACATAACACGTACTGTGTGATGAGGCATCAATGAG 2902
2341 ACTCACAATTTCTCTCTCTGTGAAATTTGCAATGGCTCTCTAGAGTACTTTGATCTCAAC 2400
2303 ACTCACAATTTCTCTCTCTGTGAAATTTGCAATGGCTCTCTAGAGTACTTTGATCTCAAC 2562
2401 ACAGACCCCTACAGAGCTGATGAATGACGTGAACACACTGGACAGGAGTGTCTCAACAG 2460
2363 ACAGACCCCTACAGAGCTGATGAATGACGTGAACACACTGGACAGGAGTGTCTCAACAG 3022
2461 CTACAGCTACAGCTCATGAGAGCTGAGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 2520
3023 CTACAGCTACAGCTCATGAGAGCTGAGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 3082
2521 ACTCGAAACATGACCTGGGACCTTAAAGATGAGAGGAGCTATGAGCAATACAGCGAGTTT 2580
3083 ACTCGAAACATGACCTG-----GATGAGAGAGCTATGAGCAATACAGCGAGTTT 3133
2581 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
3134 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193
2641 GAAGGCTGGGAAAGTTAAGAAACAAAGAGGTGACCTCCAAACATAGAGGATCAACC 2700
3194 GAAGGCTGGGAAAGTTAAGAAACAAAGAGGTGACCTCCAAACATAGAGGATCAACC 3253
2701 TGACTGACACAGGCAATGAAGAAACCATGTGGTGTGATTTCCAGCAGACCTGTGTATTGGCC 2760
3254 TGACTGACACAGGCAATGAAGAAACCATGTGGTGTGATTTCCAGCAGACCTGTGTATTGGCC 3313
2761 AGGAGGCTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC 2820
3314 AGGAGGCTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC 3373
2821 CAGCAGAGCAGAGATAAATTTAGAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880
3374 CAGCAGAGCAGAGATAAATTTAGAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3433
2881 TACCTCACAGCTGCAACAAATGCAATTTTGTATCAAAAAGTCAACCTACCTACCTCC 2940
3434 TACCTCACAGCTGCAACAAATGCAATTTTGTATCAAAAAGTCAACCTACCTACCTCC 3493
2941 CCAGAAGCTCAACAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
3494 CCAGAAGCTCAACAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553
3001 CCCAAGGCGGAAAGTCAATGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3060
3554 CCCAAGGCGGAAAGTCAATGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3613
3061 TCCTCTTATTTCTTTGCTTTGTCAAGAGAGAACTTAAGAGCAGACAGAGGCAACGT 3120
3614 TCCTCTTATTTCTTTGCTTTGTCAAGAGAGAACTTAAGAGCAGACAGAGGCAACGT 3673
3121 GGAGAGGCTGAAAACAGTGCAGAGACGTTTGAACAATGAGTCACTAGCAGCAAAAAGAGATGA 3180
3674 GGAGAGGCTGAAAACAGTGCAGAGACGTTTGAACAATGAGTCACTAGCAGCAAAAAGAGATGA 3733
3181 CATTTACTAGCAGTATAAACCTTGGTGTCTGAGAACTGCCTTCTATATATAT 3240
3734 CATTTACTAGCAGTATAAACCTTGGTGTCTGAGAACTGCCTTCTATATATAT 3793
3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTTAATAAGAAATCCAA 3300
3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTTAATAAGAAATCCAA 3853
3301 TTTTCAGAGGTGGTGTGTAATAACGCTCTGTGGCAGTGTGTAAGAAAAA 3353
3854 TTTTCAGAGGTGGTGTGTAATAACGCTCTGTGGCAGTGTGTAAGAAAAA 3906

Db 1343 ATCAAGCGAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACG 1402
QY 841 GGGCCCATGAAGCCCATCCACATGGATTCACCAACATGCTCCAGCGGAAGCGCTTGCAG 900
Db 1403 GGGCCCATGAAGCCCATCCACATGGATTCACCAACATGCTCCAGCGGAAGCGCTTGCAG 1462
QY 901 ACCCTCATGTCGGTGGACGACTCCATGGAGACGATTTTACAACATGCTGGTGGAGCGGC 960
Db 1463 ACCCTCATGTCGGTGGACGACTCCATGGAGACGATTTTACAACATGCTGGTGGAGCGGC 1522
QY 961 GAGCTGGAACAAAGCTATACGCTATACACGCGGACCAAGGTTACACATCGGCCAGTTT 1020
Db 1523 GAGCTGGAACAAAGCTATACGCTATACACGCGGACCAAGGTTTACACATCGGCCAGTTT 1582
QY 1021 GGGCTGTGTGAAGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTTACGCTG 1080
Db 1583 GGGCTGTGTGAAGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTTACGCTG 1642
QY 1081 AGGGGCCCCAAAGCTGGAGCGGCTGTCTGAATCCCAACATGCTCTCAACATTTGACCTG 1140
Db 1643 AGGGGCCCCAAAGCTGGAGCGGCTGTCTGAATCCCAACATGCTCTCAACATTTGACCTG 1702
QY 1141 GCCCCACCATCTGGACATTTGAGGCTGACATACCTCGGATATGACGCGGAATCC 1200
Db 1703 GCCCCACCATCTGGACATTTGAGGCTGACATACCTCGGATATGACGCGGAATCC 1762
QY 1201 ATCTCAAGCTGTGGACCGGCTGTCTGAATCCCAACATGCTCTCAACATTTGACCTG 1260
Db 1763 ATCTCAAGCTGTGGACCGGCTGTCTGAATCCCAACATGCTCTCAACATTTGACCTG 1822
QY 1261 AGGGCTGGGGGACTCTTCTTGGTGGAGAGGCAAGCTGTACACAAAGAGACAT 1320
Db 1823 AGGGCTGGGGGACTCTTCTTGGTGGAGAGGCAAGCTGTACACAAAGAGACAT 1882
QY 1321 GACAAAGTGGACGCCACAGGAGGAGAACTTTCTGCCAAAGTACAGCGCTGTGAAGACCTG 1380
Db 1883 GACAAAGTGGACGCCACAGGAGGAGAACTTTCTGCCAAAGTACAGCGCTGTGAAGACCTG 1942
QY 1381 TGTCAAGCTGTGATACCAAGCGGCTGTGAGAGCTGGGACAGAAAGTGGCAAGTGTGTG 1440
Db 1943 TGTCAAGCTGTGATACCAAGCGGCTGTGAGAGCTGGGACAGAAAGTGGCAAGTGTGTG 2002
QY 1441 GAGGACCGCAGCGGGAAGCTGAGCTGATAGTGAAGGCGCCCATGCGGCTGGGCGGC 1500
Db 2003 GAGGACCGCAGCGGGAAGCTGAGCTGATAGTGAAGGCGCCCATGCGGCTGGGCGGC 2062
QY 1501 AGCAGAGCCCTCTCCAACTCGTGCCCAAGTACTACGCGGAGGCGGAGCGGCTGCACC 1560
Db 2063 AGCAGAGCCCTCTCCAACTCGTGCCCAAGTACTACGCGGAGGCGGAGCGGCTGCACC 2122
QY 1561 TGTCAAGCGGGGACTACAGCTCAGCTGCGGAGCGCGGAGGAGAACTCTTCAAGAAAG 1620
Db 2123 TGTCAAGCGGGGACTACAGCTCAGCTGCGGAGCGCGGAGGAGAACTCTTCAAGAAAG 2182
QY 1621 AAGTACAAGGCGGCTATGTCGAGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 1680
Db 2183 AAGTACAAGGCGGCTATGTCGAGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2242
QY 1681 GGCAGGCTGACACGCTAGGCTGGGTGATGCCGCCAGCCCGGAACCTTCAACAGCGG 1740
Db 2243 GGCAGGCTGACACGCTAGGCTGGGTGATGCCGCCAGCCCGGAACCTTCAACAGCGG 2302
QY 1741 CACTGGCCAGGGGCGCTGAGGACCAAGATGACAAGGATGGTGGGAGCTTCACTGGCACT 1800
Db 2303 CACTGGCCAGGGGCGCTGAGGACCAAGATGACAAGGATGGTGGGAGCTTCACTGGCACT 2362
QY 1801 GGAGGCTTCCGACTACTAGCGGCCAAACCCGATTAAGTGACATCGGTGCTACATC 1860
Db 2363 GGAGGCTTCCGACTACTAGCGGCCAAACCCGATTAAGTGACATCGGTGCTACATC 2422
QY 1861 CTAGAGAACGACAGCTCAGTGTGACTGGACCTGTACAAAGTCCCTGACGGCTGGAAA 1920

Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGTGACAGCTGTACAAGTCCCTGACGGCTGGAAA 2482
QY 1921 GACCACAAGCTGACATTCGACACGAGATTGAACCTCTGACAGAACAAAATTAAGAACCTG 1980
Db 2483 GACCACAAGCTGACATTCGACACGAGATTGAACCTCTGACAGAACAAAATTAAGAACCTG 2542
QY 1981 AGGGAATCCGAGGTCACTTGAAGAAAAGCGGCGCAGAGAAGATGTGACTGTCTCAAAATC 2040
Db 2543 AGGGAATCCGAGGTCACTTGAAGAAAAGCGGCGCAGAGAAGATGTGACTGTCTCAAAATC 2602
QY 2041 AGCTTACCACACCCAGCACAAGCGGCTCAAGCACAAGAGGCTCCAGTCTGCAATCTTTC 2100
Db 2603 AGCTTACCACACCCAGCACAAGCGGCTCAAGCACAAGAGGCTCCAGTCTGCAATCTTTC 2662
QY 2101 AGGAAGGGCTTGCAGAGAAAGCAAGGTGTGCTGTTCGGGAGCAGAAAGCGCAAGAAG 2160
Db 2663 AGGAAGGGCTTGCAGAGAAAGCAAGGTGTGCTGTTCGGGAGCAGAAAGCGCAAGAAG 2722
QY 2161 AAATCTCCGAGCTGCTCAAGCGGCTCGAAGAAACAAAGACAGCTGCAGCATGCCAGGCTC 2220
Db 2723 AAATCTCCGAGCTGCTCAAGCGGCTCGAAGAAACAAAGACAGCTGCAGCATGCCAGGCTC 2782
QY 2221 ACCTGCTTCAACCCACGACAAACAGCACTGGCAGACGGCGCTTTCTGGACACTTGGGCGCT 2280
Db 2783 ACCTGCTTCAACCCACGACAAACAGCACTGGCAGACGGCGCTTTCTGGACACTTGGGCGCT 2842
QY 2281 TTCTGTGCTGACACGCGGCAACATTAACAGTACTGTGTGATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGACACGCGGCAACATTAACAGTACTGTGTGATGAGGACCATCAATGAG 2902
QY 2341 ACTCACAATTTCTCTCTGTGAAATTTGCAACTGGCTTCTCTAGAGTACTTTGATCTCAAC 2400
Db 2903 ACTCACAATTTCTCTCTGTGAAATTTGCAACTGGCTTCTCTAGAGTACTTTGATCTCAAC 2962
QY 2401 ACAGACCCCTACAGCTGTGATGATGAGTGAACACACTGGACAGGATGTCTTCAACAG 2460
Db 2963 ACAGACCCCTACAGCTGTGATGATGAGTGAACACACTGGACAGGATGTCTTCAACAG 3022
QY 2461 CTACACGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2520
Db 3023 CTACACGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 3082
QY 2521 ACTCGAAAATGGACCTTGGGACTTTAAAGATGGAGGAAGCTATGAGCAATACAGGCACTTT 2580
Db 3083 ACTCGAAAATGGACCTG-----GATGGAGGAAGCTATGAGCAATACAGGCACTTT 3133
QY 2581 CAGGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
Db 3134 CAGGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193
QY 2641 GAAGGCTGGGAAGGTTAAGAAACAAACAGAGGTGGACCTCCAAAACATAGAGGCAATCACC 2700
Db 3194 GAAGGCTGGGAAGGTTAAGAAACAAACAGAGGTGGACCTCCAAAACATAGAGGCAATCACC 3253
QY 2701 TGACTGCAAGGCAATGAAGAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3254 TGACTGCAAGGCAATGAAGAAACCAATGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 3313
QY 2761 AGGAGGCTGAGAAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3314 AGGAGGCTGAGAAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
QY 2821 CAGCAGGAGCAGAGATTAATCTTCAGGAAGTCCATTTTGGTCCCTGCTTTTGGTTGGATTA 2880
Db 3374 CAGCAGGAGCAGAGATTAATCTTCAGGAAGTCCATTTTGGTCCCTGCTTTTGGTTGGATTA 3433
QY 2881 TACCTTCAACGCTGCACAAAATGATTTTTCGATCAAAAAGTCAACACTTAACCCCTCCC 2940
Db 3434 TACCTTCAACGCTGCACAAAATGATTTTTCGATCAAAAAGTCAACACTTAACCCCTCCC 3493
QY 2941 CCAGAGCTCAAAAAGGAAAACGGAGAGAGCAGCAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCAAAAAGGAAAACGGAGAGAGCAGCAGAGAGATTTCTTGGAAATTTCT 3553

QY	541	AACTCCCGCTTTTATTAACATAACGCTGTGTGGAACGGGGTGAAGAAAAAGCAGCGCTCC	600
Db	1103	AACTCCCGCTTTTATTAACATAACGCTGTGTGGAACGGGGTGAAGAAAAAGCAGCGCTCC	1162
QY	601	GACTACTCCAAGGATTACCTCACAGACCTCATCACCAATGACAGCGTGAAGTCTTCCGCG	660
Db	1163	GACTACTCCAAGGATTACCTCACAGACCTCATCACCAATGACAGCGTGAAGTCTTCCGCG	1222
QY	661	AGTCCAAGAAGATGATCCCGCACAGCGCCAGTCTCTCATGGTCATCAGCCATCAGCCGCC	720
Db	1223	AGTCCAAGAAGATGATCCCGCACAGCGCCAGTCTCTCATGGTCATCAGCCATCAGCCGCC	1282
QY	721	CAGGCGCTTGAGATTTCAGCCCCACAAATATTACGGCTCTTCCCAACCGCATCTCAGCAC	780
Db	1283	CAGGCGCTTGAGATTTCAGCCCCACAAATATTACGGCTCTTCCCAACCGCATCTCAGCAC	1342
QY	781	ATCACGCCGAGCTACAACATACGCGCCCAACCCCGGACAAACACTGGATCATGGCTACACG	840
Db	1343	ATCACGCCGAGCTACAACATACGCGCCCAACCCCGGACAAACACTGGATCATGGCTACACG	1402
QY	841	GGGCCCATGAAGCCCATCCACATGGAAATTCACCAATATGCTCCAGCGGAAGCGCTTGCAG	900
Db	1403	GGGCCCATGAAGCCCATCCACATGGAAATTCACCAATATGCTCCAGCGGAAGCGCTTGCAG	1462
QY	901	ACCTCATGTGCGTGGACAGCTCCATGAGACCATTTTCAACATGTGTGTTGAGACGGCG	960
Db	1463	ACCTCATGTGCGTGGACAGCTCCATGAGACCATTTTCAACATGTGTGTTGAGACGGCG	1522
QY	961	GAGCTGGACAACACGTACATCTGATACGCGCGACCAACGGTTTACACATCGGCCAGTTTT	1020
Db	1523	GAGCTGGACAACACGTACATCTGATACGCGCGACCAACGGTTTACACATCGGCCAGTTTT	1582
QY	1021	GGCTGTGTGAAGGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG	1080
Db	1583	GGCTGTGTGAAGGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG	1642
QY	1081	AGGGGCCCAACGTGGAAGCGCGCTGTCTGAATCCCCACATCTGCTCTCAACATGACCTG	1140
Db	1643	AGGGGCCCAACGTGGAAGCGCGCTGTCTGAATCCCCACATCTGCTCTCAACATGACCTG	1702
QY	1141	GCCCCACCATCTGGACATGTGACGGCTGGACATACCTTCCGATATGACCGGAATCC	1200
Db	1703	GCCCCACCATCTGGACATGTGACGGCTGGACATACCTTCCGATATGACCGGAATCC	1762
QY	1201	ATCTCTAAGCTGTGGACACGGACGGCCCGGTGAATCGGTTTCACTTGAATAAAGAAAGATG	1260
Db	1763	ATCTCTAAGCTGTGGACACGGACGGCCCGGTGAATCGGTTTCACTTGAATAAAGAAAGATG	1822
QY	1261	AGGTCGTGGCGGACTCTCTTCTTGTGGAGAGGCAAGCTGTCTACAACAAGAGACAAT	1320
Db	1823	AGGTCGTGGCGGACTCTCTTCTTGTGGAGAGGCAAGCTGTCTACAACAAGAGACAAT	1882
QY	1321	GACAAGGTGGAACGCCCAGGAGGAGAACTTTCTGCCAAGTACCAGCGTGTGAAGGACCTG	1380
Db	1883	GACAAGGTGGAACGCCCAGGAGGAGAACTTTCTGCCAAGTACCAGCGTGTGAAGGACCTG	1942
QY	1381	TGTCAGCGTGTCTGAGTACCAAGACGGCTGTGACAGCTGGGACAGAAGTGGCAGTGTGTG	1440
Db	1943	TGTCAGCGTGTCTGAGTACCAGACGGCTGTGACAGCTGGGACAGAAGTGGCAGTGTGTG	2002
QY	1441	GAGAGCCCAACGGGAGCTGMACTGCATAAGTGCAGAGGCCCATCGCGCTGGCGCGC	1500
Db	2003	GAGAGCCCAACGGGAGCTGMACTGCATAAGTGCAGAGGCCCATCGCGCTGGCGCGC	2062
QY	1501	AGCAGGCCCTCTCCAACTCTGTCGCCCAAGTACTACGGGACGGGACCGGAGGCTGCACC	1560
Db	2063	AGCAGGCCCTCTCCAACTCTGTCGCCCAAGTACTACGGGACGGGACCGGAGGCTGCACC	2122
QY	1561	TGTGACACGGGGACTACAAGCTCAGGCTGGCCGGACGCCGGAAAAAATCTTTCAAGAAG	1620
Db	2123	TGTGACACGGGGACTACAAGCTCAGGCTGGCCGGACGCCGGAAAAAATCTTTCAAGAAG	2182

1621	QY	AGTACAAGCGCAGCTATGTCTCGCAGTCTCGCTCATCTCCGTCAGTGGCCATCGAGGTGGAC	1680
2183	Db	AAATAACAAGGCCAGCTATGTCTCGCAGTCTCGCTCATCTCCGTCAGTGGCCATCGAGGTGGAC	2242
1681	QY	GGCAGGGTGTACCAAGCTAGAGCCCTGGGTGATGCGCGCCAGCCCCGAAACCTTCACCAAGCGG	1740
2243	Db	GGCAGGGTGTACCAAGCTAGAGCCCTGGGTGATGCGCGCCAGCCCCGAAACCTTCACCAAGCGG	2302
1741	QY	CACCTGCCACAGGGCCCTCTGAGGACCAAGATGACAAGATGGTGGGACATTCAGTGGCACT	1800
2303	Db	CACCTGCCACAGGGCCCTCTGAGGACCAAGATGACAAGATGGTGGGACATTCAGTGGCACT	2362
1801	QY	GGAGGGCTTCCGCGACTACTCAGCGCGCAACCCCATTAAGATGACACATCGTGCTACATC	1860
2363	Db	GGAGGGCTTCCGCGACTACTCAGCGCGCAACCCCATTAAGATGACACATCGTGCTACATC	2422
1861	QY	CTAGAGAACGACACACGTCTCAGTGTGACCTGGACCTGTACAAGTCCCTGCGAGGCTGGAAA	1920
2423	Db	CTAGAGAACGACACACGTCTCAGTGTGACCTGGACCTGTACAAGTCCCTGCGAGGCTGGAAA	2482
1921	QY	GACCAACAAGCTGCACATCGACACGAGATGTAACCCCTGCAGAACAAAAATTAAAGAACCTG	1980
2483	Db	GACCAACAAGCTGCACATCGACACGAGATGTAACCCCTGCAGAACAAAAATTAAAGAACCTG	2542
1981	QY	AGGGAAGTCCGAGGTCACTCTGAAGAAAAAGCGCCAGAGAAGATGTGACTGTCAAAAATC	2040
2543	Db	AGGGAAGTCCGAGGTCACTCTGAAGAAAAAGCGCCAGAGAAGATGTGACTGTCAAAAATC	2602
2041	QY	AGCTACACACCCAGACACAAAGGCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTC	2100
2603	Db	AGCTACACACCCAGACACAAAGGCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTC	2662
2101	QY	AGGAAAGGCGCTGCAAGAGAGGACAAAGGTGTGGCTGTGTGGCGAGAGAGACGCAAGAAG	2160
2663	Db	AGGAAAGGCGCTGCAAGAGAGGACAAAGGTGTGGCTGTGTGGCGAGAGAGACGCAAGAAG	2722
2161	QY	AAACTCGCAAGTGTCTAAGCGCCTTGCGAAACAACGACAGTGCAGCATGCGAGGCCTC	2220
2723	Db	AAACTCGCAAGTGTCTAAGCGCCTTGCGAAACAACGACAGTGCAGCATGCGAGGCCTC	2782
2221	QY	ACGTGCTTCAACCAACGACACCTGCGACAGCTGCGACAGCGCGCTTTTCTTGACACACTGGGGCCT	2280
2783	Db	ACGTGCTTCAACCAACGACACCTGCGACAGCTGCGACAGCGCGCTTTTCTTGACACACTGGGGCCT	2842
2281	QY	TTCTGTGCTCTGACACGAGCGCCAAACAATAACAGTACTGGTGCATGAGGACCATCAATGAG	2340
2843	Db	TTCTGTGCTCTGACACGAGCGCCAAACAATAACAGTACTGGTGCATGAGGACCATCAATGAG	2902
2341	QY	ACTCACAAATTCCTCTCTGTGAAATTTGCAATTCGGCTTCCTAGAGTACTTTGATCTCAAC	2400
2903	Db	ACTCACAAATTCCTCTCTGTGAAATTTGCAATTCGGCTTCCTAGAGTACTTTGATCTCAAC	2962
2401	QY	ACAGACCCCTTACCAGCTGATGAATGCAGTGAACAACAACCTGGACAGGGATGTCCTCAACCG	2460
2963	Db	ACAGACCCCTTACCAGCTGATGAATGCAGTGAACAACAACCTGGACAGGGATGTCCTCAACCG	3022
2461	QY	CTACACGTAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACCAACTGTGG	2520
3023	Db	CTACACGTAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACCAACTGTGG	3082
2521	QY	ACTTCGAAACATGGACCTGGGACTTAAAGATGCGAGGAAGCTATGAGCAATACAGGCAGTTT	2580
3083	Db	ACTTCGAAACATGGACCTGGGACTTAAAGATGCGAGGAAGCTATGAGCAATACAGGCAGTTT	3133
2581	QY	CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACCAACTGTGG	2640
3134	Db	CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACCAACTGTGG	3193
2641	QY	GAGGGCTGGGAAGGTTAAGAAAACAACAGAGGTGACCTTCCAAAACATAGAGGCATCACC	2700
3194	Db	GAGGGCTGGGAAGGTTAAGAAAACAACAGAGGTGACCTTCCAAAACATAGAGGCATCACC	3253
2701	QY	TGACTGCACAGGCCAATGAAAAACCATGTGGGTGATTTTCAGACGACCTGTGCTATTTGGCC	2760

Db 3254 TGACTGCAGAGCAATGAAAAACCATGTGGTGTATTCAGCAGACCTGTCTATTGGCC 3313
QY 2761 AGGAGCCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATCTTGAGGATTAAC 2820
Db 3314 AGGAGCCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATCTTGAGGATTAAC 3373
QY 2821 CAGCAGAGCAGAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTA 2880
Db 3374 CAGCAGAGCAGAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTA 3433
QY 2881 TACCTCACCAGCTGCACAAATGCAATTTTTCGTATCAAAAAGTCACTACCTCTCC 2940
Db 3434 TACCTCACCAGCTGCACAAATGCAATTTTTCGTATCAAAAAGTCACTACCTCTCC 3493
QY 2941 CCAGAAAGCTCACAAGAGAAAACGAGAGAGCGAGAGAGATTTCCCTTGGAAATTTCT 3000
Db 3494 CCAGAAAGCTCACAAGAGAAAACGAGAGAGCGAGAGAGATTTCCCTTGGAAATTTCT 3553
QY 3001 CCCAAGGGGAAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGAGTCTCTGTTCTAAA 3060
Db 3554 CCCAAGGGGAAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGAGTCTCTGTTCTAAA 3613
QY 3061 TCCTCTATTCTTTTGGTTTGTCAAAAGAGAACTTAAGAGAGCAGAGAGCAAGT 3120
Db 3614 TCCTCTATTCTTTTGGTTTGTCAAAAGAGAACTTAAGAGAGCAGAGAGCAAGT 3673
QY 3121 GGAGAGCTGAAAACAGTGCAGAGAGCTTTGCAATGAGTCAAGTCAACAAAGAGATGA 3180
Db 3674 GGAGAGCTGAAAACAGTGCAGAGAGCTTTGCAATGAGTCAAGTCAACAAAGAGATGA 3733
QY 3181 CATTTACCTAGCACTATAACCTGCTTCCCTCTGAAGAACTGCTTCATTGTATATAT 3240
Db 3734 CATTTACCTAGCACTATAACCTGCTTCCCTCTGAAGAACTGCTTCATTGTATATAT 3793
QY 3241 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAAGAAATCCCAA 3300
Db 3794 GTGACTATTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAAGAAATCCCAA 3853
QY 3301 TTTTGAAGAGTGTGTGTGTCATTAACGCTCTGTGGCCAGTGTAAAGAAAA 3353
Db 3854 TTTTGAAGAGTGTGTGTGTCATTAACGCTCTGTGGCCAGTGTAAAGAAAA 3906

RESULT 11

ABL95597
ID ABL95597 standard; cDNA; 3906 BP.
XX ABL95597;
XX AC
XX 19-JUL-2002 (first entry)
XX Human angiogenesis related cDNA PRO1120 SEQ ID NO: 73.
XX DE
XX Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;
KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;
KW cardiant; cytostatic; antiangiogenic; hypotensive; vulnery;
KW antiarteriosclerotic; gene; ss.
OS Homo sapiens.
XX
XX WO200208284-A2.
XX
XX 31-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-US21735.
XX
XX 20-JUL-2000; 2000US-219556P.
PR 25-JUL-2000; 2000US-220624P.
PR 25-JUL-2000; 2000US-220664P.
PR 28-JUL-2000; 2000WO-US20710.
PR 02-AUG-2000; 2000US-222695P.
PR 17-AUG-2000; 2000US-0643657.

23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 07-SEP-2000; 2000US-230978P.
PR 15-SEP-2000; 2000US-000000P.
PR 18-SEP-2000; 2000US-064610.
PR 18-SEP-2000; 2000US-066530.
PR 24-OCT-2000; 2000US-242923P.
PR 08-NOV-2000; 2000US-0709238.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000US-0747259.
PR 20-DEC-2000; 2000WO-US34956.
PR 22-JAN-2001; 2001US-0767609.
PR 28-FEB-2001; 2001US-0796498.
PR 01-MAR-2001; 2001WO-US06520.
PR 09-MAR-2001; 2001US-0802706.
PR 14-MAR-2001; 2001US-0808889.
PR 22-MAR-2001; 2001US-0816744.
PR 05-APR-2001; 2001US-0828366.
PR 10-MAY-2001; 2001US-0854208.
PR 10-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001US-0866028.
PR 25-MAY-2001; 2001US-0866034.
PR 25-MAY-2001; 2001WO-US17092.
PR 30-MAY-2001; 2001US-0870574.
PR 30-MAY-2001; 2001WO-US17443.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 28-JUN-2001; 2001WO-US00000.
XX (GETH) GENENTECH INC.
PA (BAKE) BAKER K P.
PA (FERR) FERRARA N.
PA (GERB) GERBER H.
PA (GERR) GERRITSEN M E.
PA (GODD) GODDARD A. E.
PA (GODO) GODOWSKI P. J.
PA (GURN) GURNEY A. L.
PA (HILL) HILLAN K J.
PA (MARS) MARSTERS S. A.
PA (PANJ) PAN J.
PA (PAON) PAONI N. F.
PA (STEP) STEPHAN J. F.
PA (WATA) WATANABE C. K.
PA (WILL) WILLIAMS P. M.
PA (WOOD) WOOD W. I.
XX
PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;
XX
XX WPI; 2002-171999/22.
DR P-PSDB; ABB95459.
XX
XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,
XX useful in diagnosis and treatment of cardiovascular (e.g. myocardial
XX infarction), endothelial or angiogenic disorders in a mammal -
XX
XX Claim 1; Fig 73; 567pp; English.
XX
XX The present invention provides the protein and coding sequences of human
XX PRO proteins. These are useful for treating or diagnosing a
XX cardiovascular, endothelial or angiogenic disorder, including cardiac
XX hypertrophy, trauma, cancer, age-related macular degeneration,
XX atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,
XX angina, myocardial infarctions, thrombopilebitis, lymphangitis, tumour
XX angiogenesis (such as breast carcinoma and liver carcinoma) and wound
XX healing. The present sequence is a coding sequence of the invention.
XX
XX Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;

Db 2663 AGAAGGGCTTCAAGAGAGGACAAAGGTGTGGCTTTGGCGAGCAGAAAGCGCAAGAG 2722
Qy 2161 AAATCCCGCAAGCTGCTCAAGCGCTGTCAGAAACAGACAGTCAGATGCCAGGCTC 2220
Db 2723 AAATCCCGCAAGCTGCTCAAGCGCTGTCAGAAACAGACAGTCAGATGCCAGGCTC 2782
Qy 2221 ACGTGTTCACCCAGCAACACAGACATCGGAGAGCGGCTTTCTTGGACATCGGGCCCT 2280
Db 2783 ACGTGTTCACCCAGCAACACAGACATCGGAGAGCGGCTTTCTTGGACATCGGGCCCT 2842
Qy 2281 TTCTGTGCTGCACCCAGCGCAACAAATTAACAGCTGCTGTCAGTGCAGGATGCTCAAGTGA 2340
Db 2843 TTCTGTGCTGCACCCAGCGCAACAAATTAACAGCTGCTGTCAGTGCAGGATGCTCAAGTGA 2902
Qy 2341 ACTCAAAATTTCTCTTCTGTAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2400
Db 2903 ACTCAAAATTTCTCTTCTGTAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2962
Qy 2401 ACAGACCCCTACCGCTGATGAATGCACTGACATGACAGGATGCTTCAACAG 2460
Db 2963 ACAGACCCCTACCGCTGATGAATGCACTGACATGACAGGATGCTTCAACAG 3022
Qy 2461 CTACACGTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 2520
Db 3023 CTACACGTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 3082
Qy 2521 ACTCGAAACATGGACCTGGGACTTAAAGATGAGGAGAGCTATGAGCAATACAGGAGTTT 2580
Db 3083 ACTCGAAACATGGACCTG-----GATGAGGAGAGCTATGAGCAATACAGGAGTTT 3133
Qy 2581 CAGCGTCGAAAGTGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640
Db 3134 CAGCGTCGAAAGTGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 3193
Qy 2641 GAAGCTGGGAAGTTTAAAGAAACAAAGAGTGAAGCTTCCAAACATAGAGGATCAAC 2700
Db 3194 GAAGCTGGGAAGTTTAAAGAAACAAAGAGTGAAGCTTCCAAACATAGAGGATCAAC 3253
Qy 2701 TGACTGCAACGCAATGAACCAATGCTGGTGTATTTCCAGACAGCTGCTATTGGCC 2760
Db 3254 TGACTGCAACGCAATGAACCAATGCTGGTGTATTTCCAGACAGCTGCTATTGGCC 3313
Qy 2761 AGGAGGCTTGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGAGGATAAC 2820
Db 3314 AGGAGGCTTGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGAGGATAAC 3373
Qy 2821 CAGAGGAGCAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 2880
Db 3374 CAGAGGAGCAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 3433
Qy 2881 TACCTCACCGCTGCACAAATGCTTTTTCGTATCAAAAGTCAACCACTAACCCCTCCC 2940
Db 3434 TACCTCACCGCTGCACAAATGCTTTTTCGTATCAAAAGTCAACCACTAACCCCTCCC 3493
Qy 2941 CCAGAGCTCAAAAGGAAACCGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCAAAAGGAAACCGAGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3053
Qy 3001 CCCAGGCGGAAGTCAATGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTTAAA 3060
Db 3554 CCCAGGCGGAAGTCAATGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTTAAA 3613
Qy 3061 TCCTCTTATTTCTTTTGGTTTGTCAAAAGAGGAATTAAGAGCAGGACAGAGGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTTGGTTTGTCAAAAGAGGAATTAAGAGCAGGACAGAGGCAACGT 3673
Qy 3121 CGAGAGCTGAAACAGTGCAGAGAGCTTTCACATAGTCACTAGTAGCAAAAGAGATGA 3180
Db 3674 CGAGAGCTGAAACAGTGCAGAGAGCTTTCACATAGTCACTAGTAGCAAAAGAGATGA 3733
Qy 3181 CATTACCTAGCACTATAAACCTGGTGGCTCTCAAGAACTGCCTTCATTGTATATAT 3240
Db 3734 CATTACCTAGCACTATAAACCTGGTGGCTCTCAAGAACTGCCTTCATTGTATATAT 3793

Qy 3241 GTGACTATTATCATGTATCAATCAATGGAACCTTTTAGGGAACCTTAATAAGAAATCCCAA 3300
Db 3794 GTGACTATTATCATGTATCAATCAATGGAACCTTTTAGGGAACCTTAATAAGAAATCCCAA 3853
Qy 3301 TTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGCGCAGTGTAAAGAAAAA 3353
Db 3854 TTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGCGCAGTGTAAAGAAAAA 3906
RESULT 12
ABL88108
ID ABL88108 standard; cDNA; 3906 BP.
XX
AC ABL88108;
XX
DT 16-MAY-2002 (first entry)
XX
DE Human PRO1120 cDNA sequence SEQ ID NO:73.
XX
KW Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;
vulnerary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;
gene therapy; cardiovascular disorder; endothelial disorder; cancer;
angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;
age-related macular degeneration; arterial restenosis; angina;
rheumatoid arthritis; myocardial infarction; thrombophlebitis;
KW lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;
wound healing; chromosome mapping; gene mapping; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200200690-A2.
XX
PD 03-JAN-2002.
XX
PF 20-JUN-2001; 2001WO-US19692.
XX
PR 23-JUN-2000; 2000US-213637P.
PR 20-JUL-2000; 2000US-219556P.
PR 25-JUL-2000; 2000US-220824P.
PR 28-JUL-2000; 2000US-220664P.
PR 02-AUG-2000; 2000WO-US20710.
PR 17-AUG-2000; 2000US-222695P.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 07-SEP-2000; 2000US-230978P.
PR 18-SEP-2000; 2000US-0664610.
PR 18-SEP-2000; 2000US-0665350.
PR 24-OCT-2000; 2000US-242922P.
PR 08-NOV-2000; 2000US-0709238.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000US-0747259.
PR 20-DEC-2000; 2000WO-US34956.
PR 22-JAN-2001; 2001US-0767609.
PR 28-FEB-2001; 2001US-0796498.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-MAR-2001; 2001WO-US08666.
PR 09-MAR-2001; 2001US-0802706.
PR 14-MAR-2001; 2001US-0808689.
PR 22-MAR-2001; 2001US-0816744.
PR 05-APR-2001; 2001US-0828366.
PR 10-MAY-2001; 2001US-0854208.
PR 10-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001US-0866028.
PR 25-MAY-2001; 2001US-0866034.
PR 25-MAY-2001; 2001WO-US17092.
PR 30-MAY-2001; 2001US-0870574.
PR 30-MAY-2001; 2001WO-US17443.
PR 01-JUN-2001; 2001WO-US17800.
XX

(GETH) GENENTECH INC.

Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A; Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF; Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;

WPI: 2002-090516/12.

WPI; Z002-050516/
P-PSDB: ABB84853.

One hundred and eighty seven nucleic acids encoding PRO polypeptides, useful in diagnosis and treatment of cardiovascular (e.g. myocardial

Claim 2; Fig 73; 565pp; English.

AB188072 to AB188258 encode the PRO proteins given in AB84817 to AB85003. The PRO proteins and polynucleotides have cardiac, antitumorigenic, hypotensive, vulnary and antihypertensive activities, and can be used in gene therapy. The PRO polynucleotides, proteins, agonists and antagonists are useful for treating or diagnosing a cardiovascular, endothelial or angiogenic disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-related macular degeneration, atherosclerosis, hypertension, arterial retinosis, rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast carcinoma and liver carcinoma) and wound healing. The PRO polynucleotides have applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping. AB188259 to AB188267 represent primers and probes used in the exemplification of the present invention.

Sequence 3906 BP: 1047 A: 1089 C: 1041 G: 729 T: 0 other:

Very Match 98.8%; Score 3323.4; DB 24; Length 3906;

but Local Similarity 99.7%; Pred. No. 0;

Local similarity	99.7%	Freq. no.	0;
Mismatches 3343;	Conservative	0;	Mismatches
		1;	Indels
		9;	Gaps

1 GACTCCGGATCCAAAAGAGACACAGATCAGCGAAAAAAGAAAGATGGCCCCCGGAGC 60
563 GACTCCCGCATCCAAAAGAGACACAGATCAGCGAAAAAAGAAAGATGGCCCCCGGAGC 622
61 CTGCTGCTGTGCTGTGTCGCCAACTGTGTTCTCCCTGCTGGGTGGAAAGCTCGGCCTTC 120
623 CTGCTGCTGTGCTGTGTCGCCAACTGTGTTCTCCCTGCTGGGTGGAAAGCTCGGCCTTC 682
121 CTGTCGACACACCGCTGAAAGGCAAGTTTTCAGAGGGAACCGGAGGAACATTCGCCGCCAAC 180
683 CTGTCGACACACCGCTGAAAGGCAAGTTTTCAGAGGGAACCGGAGGAACATTCGCCGCCAAC 742
181 ATCATCTGCTGTGTGACGGAACACAGGATGTGAGAGTGTGGTGTCCATTCAGAGTGAATGAAC 240
743 ATCATCTGCTGTGTGACGGAACACAGGATGTGAGAGTGTGGTGTCCATTCAGAGTGAATGAAC 802
241 AAGACCCGGCGGATCATGTAGAGAGAGGGCGGGCGACATTCATCAACGGCTTCGTGACCCACA 300
803 AAGACCCGGCGGATCATGTAGAGAGAGGGCGGGCGACATTCATCAACGGCTTCGTGACCCACA 862
301 CCCATGTGCTGCCCTTCAGGCTCCTCCATCTCTCATCTGGCAAGTACGTGCACAAACACCAAC 360
863 CCCATGTGCTGCCCTTCAGGCTCCTCCATCTCTCATCTGGCAAGTACGTGCACAAACACCAAC 922
361 ACCTACACCAACATGAGAACTGCTCTCTCGCCCTCCTGGCAGGCAACAGCAGAGAGCGCG 420
923 ACCTACACCAACATGAGAACTGCTCTCTCGCCCTCCTGGCAGGCAACAGCAGAGAGCGCG 982
421 ACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGAACAGCTTTCTTCGGGGAAGTATCTTT 480
983 ACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGAACAGCTTTCTTCGGGGAAGTATCTTT 1042
481 AATGAATACAAACGGCTCCTCATGTGCGACCCCGCTGGGAAGAGTGGGTGCGACTCCTTTAA 540
1043 AATGAATACAAACGGCTCCTCATGTGCGACCCCGCTGGGAAGAGTGGGTGCGACTCCTTTAA 1102
541 AACTCCCGCTTTTATAACTACACGCTGTGTGGAACCGGGGTGAAAGAAAAAGACGGCTCC 600

Db 2183 AAGTACAAAGGCCAGCTATGTCCGAGTGTCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2242
QY GCGAGGCTGTACCAAGTGGCTGGGTGATGCGCGCCAGCCCGAAGACCTTCAACCAAGCGG 1740
Db 2243 GGCAGGCTGTACCAAGTGGCTGGGTGATGCGCGCCAGCCCGAAGACCTTCAACCAAGCGG 2302
QY 1741 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACCAAGATGGTGGGACTTTCAGTGGCACT 1800
Db 2303 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACCAAGATGGTGGGACTTTCAGTGGCACT 2362
QY 1801 GGAGGCTTCCGACTACTCAGCGCGCAACCCCAATTAAGTGCACATCGGTGCTACATC 1860
Db 2363 GGAGGCTTCCGACTACTCAGCGCGCAACCCCAATTAAGTGCACATCGGTGCTACATC 2422
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTACAAAGTCCCTGCGAGGCTGAAA 1920
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTACAAAGTCCCTGCGAGGCTGAAA 2482
QY 1921 GACACAGCTGCACATCGACCGACAGATTTGAACCTCTGCAGACACAAATTAAGACCTG 1980
Db 2483 GACACAGCTGCACATCGACCGACAGATTTGAACCTCTGCAGACACAAATTAAGACCTG 2542
QY 1981 AGGAAGTCCGAGTCCAGTGTGACCTGACCTGACCTGTACAAAGTCCCTGCGAGGCTGAAA 2040
Db 2543 AGGAAGTCCGAGTCCAGTGTGACCTGACCTGACCTGTACAAAGTCCCTGCGAGGCTGAAA 2602
QY 2041 AGTACACACCCAGCACAAGCGCGCTTACAGACAGAGGCTTCCAGTGTGATTCCTTTT 2100
Db 2603 AGTACACACCCAGCACAAGCGCGCTTACAGACAGAGGCTTCCAGTGTGATTCCTTTT 2662
QY 2101 AGGAAGGCTTCAAGAGAGGACCAAGTGTGGCTGTTCGGGAGCAGAGGCAAGCAAG 2160
Db 2663 AGGAAGGCTTCAAGAGAGGACCAAGTGTGGCTGTTCGGGAGCAGAGGCAAGCAAG 2722
QY 2161 AAATCCGCAAGCTCTCAAGCGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2220
Db 2723 AAATCCGCAAGCTCTCAAGCGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2782
QY 2221 ACFTGCTTCAACACAGCAACCAAGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2280
Db 2783 ACFTGCTTCAACACAGCAACCAAGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2842
QY 2281 TTCTGTGCTGCACAGCGCCCAACAATACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
Db 2843 TTCTGTGCTGCACAGCGCCCAACAATACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2902
QY 2341 ACTCAAAATTTCTCTCTCTGTAATTTGCAATTTGCAATTTGCAATTTGCAATTTG 2400
Db 2903 ACTCAAAATTTCTCTCTCTGTAATTTGCAATTTGCAATTTGCAATTTGCAATTTG 2962
QY 2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGACAGGAGTGTCTCAAC 2460
Db 2963 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGACAGGAGTGTCTCAAC 3022
QY 2461 CTACAGTACAGCTCAGTACAGCTGAGAGCTGCAAGGCTTACAGAGTGTACCCCGG 2520
Db 3023 CTACAGTACAGCTCAGTACAGCTGAGAGCTGCAAGGCTTACAGAGTGTACCCCGG 3082
QY 2521 ACTGAAACATGGACCTGGGACTTAAAGTGGAGAGCTATGAGCAATACAGGAGTTT 2580
Db 3083 ACTGAAACATGGACCTG-----GATGGAGAGCTATGAGCAATACAGGAGTTT 3133
QY 2581 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCAGCTGGGACAACTGTGG 2640
Db 3134 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCAGCTGGGACAACTGTGG 3193
QY 2641 GAGGCTGGGAGCTTAAAGAACACAGAGTGTGACCTTCCAAACATAGAGCACTCACC 2700
Db 3194 GAGGCTGGGAGCTTAAAGAACACAGAGTGTGACCTTCCAAACATAGAGCACTCACC 3253
QY 2701 TGACTGCACAGGCAATGAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3254 TGACTGCACAGGCAATGAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313

QY 2761 AGGAGGCTGAGAAAGCAAGCAAGCAAGCTCTCAGTCAACATGACAGATTTCTGGAGATAAC 2820
Db 3314 AGGAGGCTGAGAAAGCAAGCAAGCAAGCTCTCAGTCAACATGACAGATTTCTGGAGATAAC 3373
QY 2821 CAGCAGGAGCAGAGATAAATCTTCAGAAAGTCAATTTTGGCCCTGCTTTGCTTTGGATTA 2880
Db 3374 CAGCAGGAGCAGAGATAAATCTTCAGAAAGTCAATTTTGGCCCTGCTTTGCTTTGGATTA 3433
QY 2881 TACCTCACAGCTGCACAAAATGCAATTTTTCGATCAAAAAGTCAACACTAACCCCTCCC 2940
Db 3434 TACCTCACAGCTGCACAAAATGCAATTTTTCGATCAAAAAGTCAACACTAACCCCTCCC 3493
QY 2941 CCAGAAGCTCAAAAAGGAAAACGGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAAGCTCAAAAAGGAAAACGGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553
QY 3001 CCCAAGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3060
Db 3554 CCCAAGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3613
QY 3061 TCCTCTTATTTCTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGACAGGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGACAGGCAACGT 3673
QY 3121 GGAGAGGCTGAAACAGTGCAGAGACGTTTTCACATGATGATGATGATGATGATGATGATG 3180
Db 3674 GGAGAGGCTGAAACAGTGCAGAGACGTTTTCACATGATGATGATGATGATGATGATGATG 3733
QY 3181 CATTTACCTAGCAGTATATAAACCCTGTTGCTCTGAGAAACTGCTCTTATTGATATAT 3240
Db 3734 CATTTACCTAGCAGTATATAAACCCTGTTGCTCTGAGAAACTGCTCTTATTGATATAT 3793
QY 3241 GTGACTATTTACATGATTAATCAACATGGGAACTTTTAGGGGAACTTATAGAAATCCCAA 3300
Db 3794 GTGACTATTTACATGATTAATCAACATGGGAACTTTTAGGGGAACTTATAGAAATCCCAA 3853
QY 3301 TTTTCAGAGTGGTGTCTCAATTAAGCCTCTGTGGCCAGTGTAAAGAAAAA 3353
Db 3854 TTTTCAGAGTGGTGTCTCAATTAAGCCTCTGTGGCCAGTGTAAAGAAAAA 3906

RESULT 13

ACA03612

ID ACA03612 standard; cDNA; 3906 BP.

XX ACA03612;

XX XX

DT 23-MAY-2003 (first entry)

XX cDNA encoding human PRO polypeptide #10.

XX Human; PRO polypeptide; secreted and transmembrane protein;
XX tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
XX differentiation; chondrocyte; tumour; genetic disorder;
XX cytosolic; gene; ss.

OS Homo sapiens.

XX US2003036180-A1.

PN 20-FEB-2003.

XX 09-MAY-2002; 2002US-0143114.

PR 31-MAR-1997; 97WO-US05230.

PR 12-JUN-1998; 98WO-US12456.

PR 14-JUL-1998; 98WO-US14552.

PR 28-AUG-1998; 98WO-US17888.

PR 10-SEP-1998; 98WO-US18824.

PR 14-SEP-1998; 98WO-US19093.

PR 14-SEP-1998; 98WO-US19094.

PR 14-SEP-1998; 98WO-US19177.

181 ATCATCTGGTGTGCTGACGGACGACAGGATGTGAGCTGGGTTCCATGAGTGATGAAC 240
182 AGGCTGTGGGGGACTCTCTTCTGGTGAGAGAGGCAAGCTGCTACACAGAGAGCAAT 1882
183 GACAAAGTGGAGCGCCAGAGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAGGACCTG 1380
184 GACAAAGTGGAGCGCCAGAGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAGGACCTG 1942
185 TGTGACGCTGCTGAGTACAGAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGAGTGTGTG 1440
186 TGTGACGCTGCTGAGTACAGAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGAGTGTGTG 2002
187 GAGGACGCGACCGGGGAGCTGAAGCTGCAATAGTGCAGGGGCCCATGCGGCTGGGCGGC 1500
188 GAGGACGCGACCGGGGAGCTGAAGCTGCAATAGTGCAGGGGCCCATGCGGCTGGGCGGC 2062
189 AGCAGAGCGCTCTCCAACTCGTGCCCAAGTACTACGGGACAGGCGAGCGGCTGCACC 1560
190 AGCAGAGCGCTCTCCAACTCGTGCCCAAGTACTACGGGACAGGCGAGCGGCTGCACC 2122
191 AGCAGAGCGCTCTCCAACTCGTGCCCAAGTACTACGGGACAGGCGAGCGGCTGCACC 1620
192 TGTGACAGCGGGGACTACAAAGCTCAGCTGGCGGAGCGCGGAGGAACTCTTCAAGAG 2182
193 TGTGACAGCGGGGACTACAAAGCTCAGCTGGCGGAGCGCGGAGGAACTCTTCAAGAG 2182
194 AAGTACAGCGGCGAGCTATGTCGCGAGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 1680
195 AAGTACAGCGGCGAGCTATGTCGCGAGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2242
196 GGAGAGGCTGTACCAAGTAGGCTGGTGATGGCGCCAGCGCCGAACTCTCAAGAGCGG 1740
197 GGAGAGGCTGTACCAAGTAGGCTGGTGATGGCGCCAGCGCCGAACTCTCAAGAGCGG 2302
198 CACTGGCGAGGCGCCCTGAGGACCAAGATGACAGGATGGTGGGAGCTTCAGTGGGACT 1800
199 CACTGGCGAGGCGCCCTGAGGACCAAGATGACAGGATGGTGGGAGCTTCAGTGGGACT 2362
200 GGAGGCGCTTCCGAGTACTCAGCGCGCAACCCCTTAAAGTGCACATCGGTGTACATC 1860
201 GGAGGCGCTTCCGAGTACTCAGCGCGCAACCCCTTAAAGTGCACATCGGTGTACATC 2422
202 CTAGAGAAAGACACAGTGTGAGCTGAGCTGTGACAGTGTGACAGTGTGACAGTGTGAC 1920
203 CTAGAGAAAGACACAGTGTGAGCTGAGCTGTGACAGTGTGACAGTGTGACAGTGTGAC 2482
204 GACCAAGCTGACATGACCAAGAGTGTGAAACCCCTGACAGAACTTAAAGTGTGACAGT 1980
205 GACCAAGCTGACATGACCAAGAGTGTGAAACCCCTGACAGAACTTAAAGTGTGACAGT 2542
206 AGGAAAGTCCGAGGTCACCTGAAAGAAAGCGGCGGAGAGATGTGACTGTGACAAATC 2040
207 AGGAAAGTCCGAGGTCACCTGAAAGAAAGCGGCGGAGAGATGTGACTGTGACAAATC 2602
208 AGCTACCAACAGCAGCAAAAGCGGCGCTCAAGCAGAGAGGCTCCAGTGTGACATCTTTTC 2100
209 AGCTACCAACAGCAGCAAAAGCGGCGCTCAAGCAGAGAGGCTCCAGTGTGACATCTTTTC 2662
210 AGGAAAGTCCGAGGTCACAGAGAGGAGCAAGTGTGCTGTGCTGGGAGCAGAGCGCAAGAG 2160
211 AGGAAAGTCCGAGGTCACAGAGAGGAGCAAGTGTGCTGTGCTGGGAGCAGAGCGCAAGAG 2722
212 AAAGTCCGAGGTCGCTCAAGCGCTGACAGCAACAGCAGCTGACAGCATGCGAGGCTTC 2220
213 AAAGTCCGAGGTCGCTCAAGCGCTGACAGCAACAGCAGCTGACAGCATGCGAGGCTTC 2782
214 ACCTGTCTTACCCACAGCAACAGCAGCTGCGGCGCTTTCTTGGACACTTGGGCGCT 2280
215 ACCTGTCTTACCCACAGCAACAGCAGCTGCGGCGCTTTCTTGGACACTTGGGCGCT 2842
216 TTCTGTGCTGTGACCAAGCGGCAACAAATGACAGTGTGCTGATGAGGACCATCAATGAG 2340
217 TTCTGTGCTGTGACCAAGCGGCAACAAATGACAGTGTGCTGATGAGGACCATCAATGAG 2902
218 ACTCACAAATTTCTTCTGTGAAATTTGCAACTTGGCTTCTTAGAGTACTTTGATCTCAAC 2400

1823 AGGCTGTGGGGGACTCTCTTCTGGTGAGAGAGGCAAGCTGCTACACAGAGAGCAAT 1882
1824 GACAAAGTGGAGCGCCAGAGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAGGACCTG 1380
1825 GACAAAGTGGAGCGCCAGAGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAGGACCTG 1942
1826 TGTGACGCTGCTGAGTACAGAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGAGTGTGTG 1440
1827 TGTGACGCTGCTGAGTACAGAGCGGCTGTGAGCAGCTGGGACAGAAAGTGGAGTGTGTG 2002
1828 GAGGACGCGACCGGGGAGCTGAAGCTGCAATAGTGCAGGGGCCCATGCGGCTGGGCGGC 1500
1829 GAGGACGCGACCGGGGAGCTGAAGCTGCAATAGTGCAGGGGCCCATGCGGCTGGGCGGC 2062
1830 AGCAGAGCGCTCTCCAACTCGTGCCCAAGTACTACGGGACAGGCGAGCGGCTGCACC 1560
1831 AGCAGAGCGCTCTCCAACTCGTGCCCAAGTACTACGGGACAGGCGAGCGGCTGCACC 2122
1832 AGCAGAGCGCTCTCCAACTCGTGCCCAAGTACTACGGGACAGGCGAGCGGCTGCACC 1620
1833 TGTGACAGCGGGGACTACAAAGCTCAGCTGGCGGAGCGCGGAGGAACTCTTCAAGAG 2182
1834 TGTGACAGCGGGGACTACAAAGCTCAGCTGGCGGAGCGCGGAGGAACTCTTCAAGAG 2182
1835 AAGTACAGCGGCGAGCTATGTCGCGAGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 1680
1836 AAGTACAGCGGCGAGCTATGTCGCGAGCTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2242
1837 GGAGAGGCTGTACCAAGTAGGCTGGTGATGGCGCCAGCGCCGAACTCTCAAGAGCGG 1740
1838 GGAGAGGCTGTACCAAGTAGGCTGGTGATGGCGCCAGCGCCGAACTCTCAAGAGCGG 2302
1839 CACTGGCGAGGCGCCCTGAGGACCAAGATGACAGGATGGTGGGAGCTTCAGTGGGACT 1800
1840 CACTGGCGAGGCGCCCTGAGGACCAAGATGACAGGATGGTGGGAGCTTCAGTGGGACT 2362
1841 GGAGGCGCTTCCGAGTACTCAGCGCGCAACCCCTTAAAGTGCACATCGGTGTACATC 1860
1842 GGAGGCGCTTCCGAGTACTCAGCGCGCAACCCCTTAAAGTGCACATCGGTGTACATC 2422
1843 CTAGAGAAAGACACAGTGTGAGCTGAGCTGTGACAGTGTGACAGTGTGACAGTGTGAC 1920
1844 CTAGAGAAAGACACAGTGTGAGCTGAGCTGTGACAGTGTGACAGTGTGACAGTGTGAC 2482
1845 GACCAAGCTGACATGACCAAGAGTGTGAAACCCCTGACAGAACTTAAAGTGTGACAGT 1980
1846 GACCAAGCTGACATGACCAAGAGTGTGAAACCCCTGACAGAACTTAAAGTGTGACAGT 2542
1847 AGGAAAGTCCGAGGTCACCTGAAAGAAAGCGGCGGAGAGATGTGACTGTGACAAATC 2040
1848 AGGAAAGTCCGAGGTCACCTGAAAGAAAGCGGCGGAGAGATGTGACTGTGACAAATC 2602
1849 AGCTACCAACAGCAGCAAAAGCGGCGCTCAAGCAGAGAGGCTCCAGTGTGACATCTTTTC 2100
1850 AGCTACCAACAGCAGCAAAAGCGGCGCTCAAGCAGAGAGGCTCCAGTGTGACATCTTTTC 2662
1851 AGGAAAGTCCGAGGTCACAGAGAGGAGCAAGTGTGCTGTGCTGGGAGCAGAGCGCAAGAG 2160
1852 AGGAAAGTCCGAGGTCACAGAGAGGAGCAAGTGTGCTGTGCTGGGAGCAGAGCGCAAGAG 2722
1853 AAAGTCCGAGGTCGCTCAAGCGCTGACAGCAACAGCAGCTGACAGCATGCGAGGCTTC 2220
1854 AAAGTCCGAGGTCGCTCAAGCGCTGACAGCAACAGCAGCTGACAGCATGCGAGGCTTC 2782
1855 ACCTGTCTTACCCACAGCAACAGCAGCTGCGGCGCTTTCTTGGACACTTGGGCGCT 2280
1856 ACCTGTCTTACCCACAGCAACAGCAGCTGCGGCGCTTTCTTGGACACTTGGGCGCT 2842
1857 TTCTGTGCTGTGACCAAGCGGCAACAAATGACAGTGTGCTGATGAGGACCATCAATGAG 2340
1858 TTCTGTGCTGTGACCAAGCGGCAACAAATGACAGTGTGCTGATGAGGACCATCAATGAG 2902
1859 ACTCACAAATTTCTTCTGTGAAATTTGCAACTTGGCTTCTTAGAGTACTTTGATCTCAAC 2400

Db 2903 ACTCACAATTTCTCTTCTGTGTAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2962
Qy 2401 ACAGAGCCCTACAGCTGATGATGAGTGAACACACTGCGACAGGATGCTCTCAACAG 2460
Db 2963 ACAGAGCCCTACAGCTGATGATGAGTGAACACACTGCGACAGGATGCTCTCAACAG 3022
Qy 2461 CTACAGCTACAGCTGATGAGTGAACAGGTTTACAAGCAGTGAACCCCGG 2520
Db 3023 CTACAGCTACAGCTGATGAGTGAACAGGTTTACAAGCAGTGAACCCCGG 3082
Qy 2521 ACTCGAATACAGTACCTGGACCTTAAAGATGAGGAGCTTATGAGCAATACAGGAGTTT 2580
Db 3083 ACTCGAATACAGTACCTGGACCTTAAAGATGAGGAGCTTATGAGCAATACAGGAGTTT 3133
Qy 2581 CAGCTCGAAGTGGCGAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
Db 3134 CAGCTCGAAGTGGCGAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193
Qy 2641 GAAGGCTGGAGGTTAAGAAACAAAGAGGTTGGACCTTCCAAACATAGAGGCTATCACC 2700
Db 3194 GAAGGCTGGAGGTTAAGAAACAAAGAGGTTGGACCTTCCAAACATAGAGGCTATCACC 3253
Qy 2701 TGACTGCACAGGCAATGAAGAAACCAATGTGGTGAATTTCCAGCAGACCTGTCTATTGGCC 2760
Db 3254 TGACTGCACAGGCAATGAAGAAACCAATGTGGTGAATTTCCAGCAGACCTGTCTATTGGCC 3313
Qy 2761 AGAGGCTGGAGAAACCAAGCAGCCTCTCAGTCAACATGACAGATTCGGAGGATTAAC 2820
Db 3314 AGAGGCTGGAGAAACCAAGCAGCCTCTCAGTCAACATGACAGATTCGGAGGATTAAC 3373
Qy 2821 CAGCAGGAGCAGAGATTAATCTCAGGAGTCCATTTTGGCCCTGCTTTTGGCTTTCGATTA 2880
Db 3374 CAGCAGGAGCAGAGATTAATCTCAGGAGTCCATTTTGGCCCTGCTTTTGGCTTTCGATTA 3433
Qy 2881 TACCTCACCAGCTGCACAAATGCAATTTTTCGTATCAAAAAGTCCACCTAACCTCTCC 2940
Db 3434 TACCTCACCAGCTGCACAAATGCAATTTTTCGTATCAAAAAGTCCACCTAACCTCTCC 3493
Qy 2941 CAGAGCTCACAAGGAAACCGAGAGAGCGAGGAGAGATTCCTTGGAAATTTCT 3000
Db 3494 CAGAGCTCACAAGGAAACCGAGAGAGCGAGGAGAGATTCCTTGGAAATTTCT 3553
Qy 3001 CCCAAGGCGCAAGTCAATGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAA 3060
Db 3554 CCCAAGGCGCAAGTCAATGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAA 3613
Qy 3061 TCCTCTTATTTCTTTGTTGTCAAAAAGAGAACTAAGAGCAGGACAGGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTGTTGTCAAAAAGAGAACTAAGAGCAGGACAGGCAACGT 3673
Qy 3121 GGAGGCTCAAAACAGTGCAGAGCGTTTGAATGAGTCAAGTACAAAGAGATGA 3180
Db 3674 GGAGGCTCAAAACAGTGCAGAGCGTTTGAATGAGTCAAGTACAAAGAGATGA 3733
Qy 3181 CATTTACCTAGCACTATAAACCTTGTGCTCTGAAGAACTGCTTCAATGTATATAT 3240
Db 3734 CATTTACCTAGCACTATAAACCTTGTGCTCTGAAGAACTGCTTCAATGTATATAT 3793
Qy 3241 GTGACTATTTACATGTAATCAATGGAATTTTGGGGAACCTTAAGAAATCCAA 3300
Db 3794 GTGACTATTTACATGTAATCAATGGAATTTTGGGGAACCTTAAGAAATCCAA 3853
Qy 3301 TTTTCAGGAGTGGTGTGTCAATAAAGCTTCTGTGGCCAGTGTAAAGAAAAA 3353
Db 3854 TTTTCAGGAGTGGTGTGTCAATAAAGCTTCTGTGGCCAGTGTAAAGAAAAA 3906

RESULT 14

ACA04033

ID ACA04033 standard; cDNA; 3906 BP.

XX

AC

XX

DT 27-MAY-2003 (first entry)
XX Human cDNA encoding a secreted/transmembrane protein, SEQ ID 19.
DE
XX Human; ss; gene; secreted protein; transmembrane protein; PRO;
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defects; premature aging; AIDS; biosensor;
KW acquired immunodeficiency syndrome; cancer; diabetic complication;
KW bioreactor; tumour.
XX
OS Homo sapiens.
XX
FN US2003032155-A1.
XX
PD 13-FEB-2003.
XX
PF 03-MAY-2002; 2002US-0137865.
XX
PR 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12352.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23059.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.

PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US202031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-MAR-2001; 2001WO-US06666.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 22-JUN-2001; 2001WO-US20116.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 20-DEC-2000; 2000US-0747259.
PR 28-FEB-2001; 2001US-0796498.
PR 09-MAR-2001; 2001US-0802706.
PR 14-MAR-2001; 2001US-0808689.
PR 22-MAR-2001; 2001US-0816744.
PR 05-APR-2001; 2001US-0828366.
PR 10-MAY-2001; 2001US-0854208.
PR 18-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001US-0860216.
PR 25-MAY-2001; 2001US-0866028.
PR 25-MAY-2001; 2001US-0866034.
PR 01-JUN-2001; 2001US-0872035.
PR 05-JUN-2001; 2001US-0874503.
PR 14-JUN-2001; 2001US-0882636.
PR 19-JUN-2001; 2001US-0886342.
PR 21-JUN-2001; 2001US-0887879.
PR 18-JUL-2001; 2001US-0908827.
PR 06-AUG-2001; 2001US-0924419.
PR 09-AUG-2001; 2001US-0927796.
PR 16-AUG-2001; 2001US-0931836.
PR 19-DEC-2001; 2001US-0028072.

(GETH) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI: 2003-331925/31.
P-PSDB; ABU66855.

XX New secreted and transmembrane nucleic acids and polypeptides,
PT designated as PRO, useful for treating inflammation, organ failure,
PT atherosclerosis, cardiac injury, infertility, birth defects, premature
PT aging, AIDS, or cancer -
XX
PS Claim 2; Fig 19; 659pp; English.

XX The invention relates to an isolated nucleic acid comprising, or which is
CC at least 80% identical to, or the full-length coding sequence of, any of
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
CC (one of 275 secreted or transmembrane proteins). The nucleic acid
CC further comprises the full-length coding sequence of the DNA deposited
CC under American Type Culture Collection (ATCC) accession number in a list
CC given in the specification. Also included are vectors and host
CC cells for producing PRO proteins, PRO fusion proteins, anti-PRO
CC antibodies, PRO extracellular domains and mature sequences, methods
CC of detecting PRO proteins, methods for stimulating the release of
CC TNF-alpha (tumour necrosis factor alpha) from human blood,
CC (and the proliferation of differentiation of chondrocyte cells, the

CC proliferation of, or gene expression in pericyte cells, the release or
CC proteoglycans from cartilage, proliferation of inner ear utricular
CC supporting cells, the proliferation of T-lymphocyte cells, the release
CC of a cytokine from peripheral blood mononuclear cells (PBMC), or the
CC proliferation of endothelial cells), a method for modulating the uptake
CC of glucose or free fatty acid (FFA) by skeletal muscle cells,
CC a method for inhibiting the binding of A-peptide to factor VIIA,
CC or the differentiation of adipocyte cells, a method for detecting the
CC presence of a tumour in a mammal and an oligonucleotide probe derived
CC from any of the nucleotide sequences cited above. The nucleic acids and
CC polypeptides are useful for treating inflammatory diseases, organ
CC failure, atherosclerosis, cardiac injury, infertility, birth defects,
CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or
CC diabetic complications. The nucleic acids are useful as hybridisation
CC probes, in chromosome and gene mapping, and in generating antisense RNA
CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,
CC biosensors or bioreactors. Both are useful in tissue typing.
CC The present sequence encodes a PRO protein of the invention.

XX
SQ Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;

Query Match 98.8%; Score 3323.4; DB 25; Length 3906;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 GACTCCGCATCCCAAAGAGAGCACCAGATCAGCAAAAAGAGAGATGGCCCCCGAGC 60
DB 563 GACTCCGCATCCCAAAGAGAGCACCAGATCAGCAAAAAGAGAGATGGCCCCCGAGC 622
QY 61 CTCGTGCTGTGCTGCTGCCCAACTGTTCCTCCCTGGTGGTGAAGCTCGGCCCTTC 120
DB 623 CTCGTGCTGTGCTGCTGCCCAACTGTTCCTCCCTGGTGGTGAAGCTCGGCCCTTC 682
QY 121 CTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAAGAGGACCGCAGAGAAATCGCCCCAAC 180
DB 683 CTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAAGAGGACCGCAGAGAAATCGCCCCAAC 742
QY 181 ATCATCTGCTGCTGACGAGCAGCAGGATGTGGAGCTGGGTTCCTCATGAGTGATGAAC 240
DB 743 ATCATCTGCTGCTGACGAGCAGCAGGATGTGGAGCTGGGTTCCTCATGAGTGATGAAC 802
QY 241 AAGACCCGCGCATCATGAGCAGGCGGGCGGCGCATTCATCAACGCTTCGTGACCAACA 300
DB 803 AAGACCCGCGCATCATGAGCAGGCGGGCGGCGCATTCATCAACGCTTCGTGACCAACA 862
QY 301 CCATGTGCTGCCCTCAGCTCCTCCATCTCTCATGCGCAAGTAGTCCACACCAAC 360
DB 863 CCATGTGCTGCCCTCAGCTCCTCCATCTCTCATGCGCAAGTAGTCCACACCAAC 922
QY 361 ACTATACCAACAATGAGAACTGCTCCTCGCCCTCCTGCGAGGCACAGCAGAGCGCG 420
DB 923 ACTATACCAACAATGAGAACTGCTCCTCGCCCTCCTGCGAGGCACAGCAGAGCGCG 982
QY 421 ACTTTGCGGTGACTCTCAATAGCACTGGCTACCGGACAGCTTTCTTGGGAAGTATCTT 480
DB 983 ACTTTGCGGTGACTCTCAATAGCACTGGCTACCGGACAGCTTTCTTGGGAAGTATCTT 1042
QY 481 AATGAATACAAGCGCTCTACGTGCCACCGGCTGGAAGGAGTGGGTGGGTCTCTTAAA 540
DB 1043 AATGAATACAAGCGCTCTACGTGCCACCGGCTGGAAGGAGTGGGTGGGTCTCTTAAA 1102
QY 541 AACTCCCGCTTTTATACTATACAGCTGTGTGGAAACGGGGTGAAGAAAGACCGGCTCC 600
DB 1103 AACTCCCGCTTTTATACTATACAGCTGTGTGGAAACGGGGTGAAGAAAGACCGGCTCC 1162
QY 601 GACTACTCCAAGGATTACTCTACAGCTCATCAATGACAGGCTGAGCTTCTTCGCG 660
DB 1163 GACTACTCCAAGGATTACTCTACAGCTCATCAATGACAGGCTGAGCTTCTTCGCG 1222
QY 661 AGCTCCAAAGAGATGTACCGCAGAGCGAGCTTCCTCATGGTTCATCAGCCATGACGCC 720
DB 1223 AGCTCCAAAGAGATGTACCGCAGAGCGAGCTTCCTCATGGTTCATCAGCCATGACGCC 1282

QY 721 CACGGCCCTGAGGATTCAGCCCCCAATATATACGCTCTTCCAAACGATCTCAGCAC 780
DB 1283 CAGGGCCCTGAGGATTCAGCCCCCAATATATACGCTCTTCCAAACGATCTCAGCAC 1342
QY 781 ATCAGCCCGAGCTCAACTACGCGCCCAACCGGACAAACACTGGATCCTGCGCTACAG 840
DB 1343 ATCAGCCCGAGCTCAACTACGCGCCCAACCGGACAAACACTGGATCCTGCGCTACAG 1402
QY 841 GGGCCCATGAAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGTTTCAG 900
DB 1403 GGGCCCATGAAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGTTTCAG 1462
QY 901 ACCCTCATGTGCTGGAGCACTCCATGAGACAGATTTACAACTGCTGGTTGAGACGGC 960
DB 1463 ACCCTCATGTGCTGGAGCACTCCATGAGACAGATTTACAACTGCTGGTTGAGACGGC 1522
QY 961 GAGCTGGGCAACACAGCTACATCGTATACACCGCGGACACAGGTTACCACTCGGCGAGTTT 1020
DB 1523 GAGCTGGGCAACACAGCTACATCGTATACACCGCGGACACAGGTTACCACTCGGCGAGTTT 1582
QY 1021 GGCCTGGTAAAGGAAATCCATGCGCATATGAGTTTGATCATCGGTTCCGTTCTACGTG 1080
DB 1583 GGCCTGGTAAAGGAAATCCATGCGCATATGAGTTTGATCATCGGTTCCGTTCTACGTG 1642
QY 1081 AGGGGCCCAACAGCTGAAAGCGGCTGTCTGAATCCCACTCTCCCAATGACCTG 1140
DB 1643 AGGGGCCCAACAGCTGAAAGCGGCTGTCTGAATCCCACTCTCCCAATGACCTG 1702
QY 1141 GCCCCACCATCTCTGGACATTTGAGGCTTGACATACCTGCGGATATGACGGGAATTC 1200
DB 1703 GCCCCACCATCTCTGGACATTTGAGGCTTGACATACCTGCGGATATGACGGGAATTC 1762
QY 1201 ATCTCTAAGCTCTGGACACCGAGCGGCGGTGATCGTTTCTCTGAAAGAGATG 1260
DB 1763 ATCTCTAAGCTCTGGACACCGAGCGGCGGTGATCGTTTCTCTGAAAGAGATG 1822
QY 1261 AGGGTCTGGCGGACTCCTTCTTGTGGAGAGGCAAGCTGTCTACACAGAGAGACAAT 1320
DB 1823 AGGGTCTGGCGGACTCCTTCTTGTGGAGAGGCAAGCTGTCTACACAGAGAGACAAT 1882
QY 1321 GACAGGTGGACCGCCAGAGAGAGAACTTTCTGCCAAGTACCAGCTGTCAAGACCTG 1380
DB 1883 GACAGGTGGACCGCCAGAGAGAGAACTTTCTGCCAAGTACCAGCTGTGTGAAGACCTG 1942
QY 1381 TGTACAGCTGTGAGTACACAGCGGCGTGTGAGCAGCTGGGACAGAACTGGCAGTGTG 1440
DB 1943 TGTACAGCTGTGAGTACACAGCGCGTGTGAGCAGCTGGGACAGAACTGGCAGTGTG 2002
QY 1441 GAGGACGCCACGGGGAAGCTGAAGCTGCATTAAGTGCAGAGGCGCCCATCGCGCTGGGCGC 1500
DB 2003 GAGGACGCCACGGGGAAGCTGAAGCTGCATTAAGTGCAGAGGCGCCCATCGCGCTGGGCGC 2062
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTAGGGCAGGCGAGGCTGCAC 1560
DB 2063 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTAGGGCAGGCGAGGCTGCAC 2122
QY 1561 TGTGACAGCGGGACTCAAGCTCAGCTTGGCGGACGCGGAAAAAACTCTTCAAGAG 1620
DB 2123 TGTGACAGCGGGACTCAAGCTCAGCTTGGCGGACGCGGAAAAAACTCTTCAAGAG 2182
QY 1621 AGTACAAGCCAGCTATCTCCGAGTCCCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1680
DB 2183 AGTACAAGCCAGCTATCTCCGAGTCCCTCCATCCGCTCAGTGGCCATCGAGTGGAC 2242
QY 1681 GGCAGGGTGTACCACTAGGCTTGGTGTATGCGGCCCGAGCCCGAAACCTCACCAAGCGG 1740
DB 2243 GGCAGGGTGTACCACTAGGCTTGGTGTATGCGGCCCGAGCCCGAAACCTCACCAAGCGG 2302
QY 1741 CACTGCGAGGGCCCTTGAGGACCAAGATGACAGGATGGTGGGACTTCTAGTGGCACT 1800
DB 2303 CACTGCGAGGGCCCTTGAGGACCAAGATGACAGGATGGTGGGACTTCTAGTGGCACT 2362
QY 1801 GGAGGCTTCCGACTACTCAGCGCGCAACCCCAATTAAGTGACACATCGGTGTACATC 1860

DB 2363 GGAGGCTTCCGACTACTCAGCGCGCAACCCCAATTAAGTGACACATCGGTGTACATC 2422
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTACAAAGTCCCTCAGGCTTGAAA 1920
DB 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTACAAAGTCCCTCAGGCTTGAAA 2482
QY 1921 GACCAACAAGTGTGACATCGACACAGATTTGAAACCTCTGCGAGAACAAATTAAGAACCTG 1980
DB 2483 GACCAACAAGTGTGACATCGACACAGATTTGAAACCTCTGCGAGAACAAATTAAGAACCTG 2542
QY 1981 AGGGAAGTCCGAGGTCACTGTAAGAAAAAGCGCCAGAAAGATGTGACTGTCACAAAATC 2040
DB 2543 AGGGAAGTCCGAGGTCACTGTAAGAAAAAGCGCCAGAAAGATGTGACTGTCACAAAATC 2602
QY 2041 AGCTACCAACCCACGACCAAAAGCGCCTCAAGCAACAGAGCTCCAGTCTGCAATCCTTTC 2100
DB 2603 AGCTACCAACCCACGACCAAAAGCGCCTCAAGCAACAGAGCTCCAGTCTGCAATCCTTTC 2662
QY 2101 AGGAAGGCTCTCAAGAGAACGACAGGTGTGGCTGTTCGCGGACGAAAGCGCAAGAG 2160
DB 2663 AGGAAGGCTCTCAAGAGAACGACAGGTGTGGCTGTTCGCGGACGAAAGCGCAAGAG 2722
QY 2161 AAACCTCCGAACTGCTCAAGCGCCTGACAGAACAAACGACAGCTGCGAGTGCAGGCTC 2220
DB 2723 AAACCTCCGAACTGCTCAAGCGCCTGACAGAACAAACGACAGCTGCGAGTGCAGGCTC 2782
QY 2221 AGTGTCTTCAACCAACGACCACTGGCAGAGCGGCTTCTGACACTGGGGCT 2280
DB 2783 AGTGTCTTCAACCAACGACCACTGGCAGAGCGGCTTCTGACACTGGGGCT 2842
QY 2281 TTCTGTGCTGACCAACGAGCCCAACAAATAACAGTACTGCTGATGAGGACCAATCAATGAG 2340
DB 2843 TTCTGTGCTGACCAACGAGCCCAACAAATAACAGTACTGCTGATGAGGACCAATCAATGAG 2902
QY 2341 ACTCAAAATTTCTCTTCTGTGAATTTGCAACTGCTTCTAGAGTACTTGTATCTCAAC 2400
DB 2903 ACTCAAAATTTCTCTTCTGTGAATTTGCAACTGCTTCTAGAGTACTTGTATCTCAAC 2962
QY 2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACACTGGAACAGGATGCTCAACCCAG 2460
DB 2963 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACACTGGAACAGGATGCTCAACCCAG 3022
QY 2461 CTACAGCTCAGCTATGAGCTGAGGAGCTGCAAGGTTTCAAGAGTGTAAACCCCGG 2520
DB 3023 CTACAGCTCAGCTATGAGCTGAGGAGCTGCAAGGTTTCAAGAGTGTAAACCCCGG 3082
QY 2521 ACTCGAAACATGACCTGGGACTTAAAGATGAGGAAGCTATGAGCAATACAGGCACTT 2580
DB 3083 ACTCGAAACATGAGCTG-----GATGAGGAGCTATGAGCAATACAGGCACTT 3133
QY 2581 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCTCTGGGACACTGTGG 2640
DB 3134 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCTCTGGGACACTGTGG 3193
QY 2641 GAAGGCTGGGAAGTTTAAAGAACACAGAGGTGGACCTCCAAACATAGAGGCACTACC 2700
DB 3194 GAAGGCTGGGAAGTTTAAAGAACACAGAGGTGGACCTCCAAACATAGAGGCACTACC 3253
QY 2701 TGACTGCACAGGCAATGAAGAAACCAATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760
DB 3254 TGACTGCACAGGCAATGAAGAAACCAATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 3313
QY 2761 AGGAGGCTGAGAAACCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
DB 3314 AGGAGGCTGAGAAACCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
QY 2821 CAGCAGGAGCAGAGATAAATCTTCAAGGAAGTCCATTTTGGCCCTGCTTTTGGTGAATTA 2880
DB 3374 CAGCAGGAGCAGAGATAAATCTTCAAGGAAGTCCATTTTGGCCCTGCTTTTGGTGAATTA 3433
QY 2881 TACTCAGCAGCTGCAAAAATGCAATTTTGTATCAAAAAGTCAACCAATCAACCTCC 2940

3434	TACCTCAGCAGCTGCACAAAATGCATTTTTTTCGTATCAAAAAGTCCACCACTAACCTCC	3493
2941	CCAGAAAGCTCACAAGAGAAAAACGGAGAGAGCGAGCGAGAGATTTTCCTTGGAAAAATTTCT	3000
3494	CCAGAAAGCTCACAAGAGAAAAACGGAGAGAGCGAGCGAGAGATTTTCCTTGGAAAAATTTCT	3553
3001	CCCAAGGGCGAAAGTCATTGGAAATTTTAAATCATATAGGGGAAAAAGCAGTCCCTGTTCTAAA	3060
3554	CCCAAGGGCGAAAGTCATTGGAAATTTTAAATCATATAGGGGAAAAAGCAGTCCCTGTTCTAAA	3613
3061	TCTCTCTATTCTTTTGGTTTGTCCAAAAAGAGAACTTAAGAAAGCAGGACAGAGGCAACCT	3120
3614	TCTCTCTATTCTTTTGGTTTGTCCAAAAAGAGAACTTAAGAAAGCAGGACAGAGGCAACCT	3673
3121	GGAGAGCTGAAACAAACAGTGCAGAGACGTTTGTACAATGAGTCAGTAGCACAAAAGAGATGA	3180
3674	GGAGAGCTGAAACAAACAGTGCAGAGACGTTTGTACAATGAGTCAGTAGCACAAAAGAGATGA	3733
3181	CATTTTACCTTAGCAGCATATAAACCTCGTTCGCCCTGTGAAGAAACGTGCCCTTCATTGTATATAT	3240
3734	CATTTTACCTTAGCAGCATATAAACCTCGTTCGCCCTGTGAAGAAACGTGCCCTTCATTGTATATAT	3793
3241	GTGACTATTTTACATGTAAATCAACATGGGAACCTTTTAGGGGAAACCTTAATAAGAAATCCCAA	3300
3794	GTGACTATTTTACATGTAAATCAACATGGGAACCTTTTAGGGGAAACCTTAATAAGAAATCCCAA	3853
3301	TTTTTCAGGAGTGGTGGTGCAATAAAGCGTCTGTGGCCAGTGTAAAGAAAAA	3353
3854	TTTTTCAGGAGTGGTGGTGCAATAAAGCGTCTGTGGCCAGTGTAAAGAAAAA	3906

RESULT 15

ABX89150
ID ABX89150 standard; cDNA; 3906 BP.

AC ABX89150;

DT 13-MAY-2003 (first entry)

DE DNA encoding novel secreted and transmembrane protein PRO1120.

Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing; cardiac insufficiency disorder; cancer; tumour; immune response; adrenal cortical capillary endothelial growth; c-fos induction; vascular endothelial growth factor inhibition; VEGF inhibition; endothelial cell growth inhibitor; T-lymphocytes stimulation; retinal neurons cell survival; rod photoreceptor cell survival; retinal disorder; retinitis pigmentosa; kidney disease; mammalian kidney mesangial cell proliferation; Berger disease; dermatitis; herpetiformis; Crohn's disease; chondrocyte proliferation; chondrocyte redifferentiation; sports injury; arthritis; gene; ss.

OS Homo sapiens.

PN US2003017563-A1.

23-JAN-2003.

07-MAY-2002; 2002US-0140808.

31-MAR-1997; 97WO-US05230.

PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.

PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.

PR 14-SEP-1998; 98WC-US19093.
PR 14-SEP-1998; 98WC-US19094.

PR 14-SEP-1998; 98WC-US19177.
PR 16-SEP-1998; 98WC-US19330.

PR	I/--SEP-1998;	98WO-US19437.
PR	07-OCT-1998;	98WO-US21141

PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992

PR 18-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 25-MAY-2001; 2001US-0866034.
 PR 01-JUN-2001; 2001US-0872035.
 PR 05-JUN-2001; 2001US-0874503.
 PR 14-JUN-2001; 2001US-0882536.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W,
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2003-148238/14.
 DR P-PSDB; ABU59660.
 XX
 PT Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
 PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
 PT are therapeutically useful for enhancing immune response and in cancer
 PT treatments -
 PT
 XX
 PS Claim 2; Fig 19; 659pp; English.
 XX
 CC The invention describes an isolated human PRO polypeptide. The PRO
 CC polypeptides are useful in detecting PRO polypeptides in a sample, in
 CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and
 CC in modulating at least one biological activity of a cell expressing a PRO
 CC polypeptide. PRO132 stimulates hypertrophy of neonatal heart and is thus
 CC useful for treating cardiac insufficiency disorders. PRO154 and PRO1186
 CC stimulate adrenal cortical capillary endothelial growth, and PRO536,
 CC PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,
 CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus
 CC useful for treating conditions or disorders where angiogenesis would be
 CC beneficial, e.g. wound healing and disorders where angiogenesis would be
 CC useful for treating cancerous tumours. PRO812 inhibits vascular
 CC endothelial growth factor (VEGF) stimulated proliferation of endothelial
 CC cells and is thus useful for inhibiting endothelial cell growth in
 CC mammals which would be beneficial in inhibiting tumour growth. PRO826,
 CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of
 CC stimulated T-lymphocytes and are therapeutically useful for enhancing
 CC immune response. PRO828, PRO826, PRO1068 or PRO132 enhance survival of
 CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of
 CC rod photoreceptor cells) and therefore are useful for treating retinal
 CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813
 CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,
 CC and therefore are useful for treating kidney disorders associated with
 CC decreased mesangial cell function such as Berger disease or other
 CC nephropathies associated with dermatitis, herpeticiformis or Crohn's
 CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
 CC proliferation and/or redifferentiation of chondrocytes in culture and
 CC are thus useful for treating sports injuries, and arthritis. This
 CC sequence encodes a novel human PRO protein.
 XX
 SQ Sequence 3906 BP; 1047 A; 1089 C; 1041 G; 729 T; 0 other;
 Query Match 98.8%; Score 3323.4; DB 25; Length 3906;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 1 GACTCCCGCATCCCAAGAACACACAGATCAGCAAAAGAAAGATGGCCCCCGGAGC 60
 563 GACTCCCGCATCCCAAGAACACACAGATCAGCAAAAGAAAGATGGCCCCCGGAGC 622
 61 CTGGTCTGTGCTGTGTCGCGCAATGTGTCTCCCTGCTGGGTGGAAGCTCGGCCTTC 120
 623 CTCGTCTGTGCTGTGTCGCGCAATGTGTCTCCCTGCTGGGTGGAAGCTCGGCCTTC 682

QY	121	CTGTGCGCACACCGCTGTAAGAGCGACGTTTTCAGAGGGAGCCGCGAGGAACATCGSCCCCAAC	180
DB	583	CTGTGCGCACACCGCTGTAAGAGCGACGTTTTCAGAGGGAGCCGCGAGGAACATCGSCCCCAAC	742
QY	181	ATCATCTGTGTGTGATGACGAGACGACAGGATGTGGAGCTGGGTTCATGAGGTGATGAAC	240
DB	743	ATCATCTGTGTGTGATGACGAGACGACAGGATGTGGAGCTGGGTTCATGAGGTGATGAAC	802
QY	241	AAGACCGCGGCATCATGAGGAGGAGGCGGCGCATTCATCAACGCTTCGTGAGCACCA	300
DB	803	AAGACCGCGGCATCATGAGGAGGAGGCGGCGCATTCATCAACGCTTCGTGAGCACCA	862
QY	301	CCCATGTGTGCTGCCCCCTCAGCTCTCTCATCTGCAAGTACGTCCACAAACACACAAC	360
DB	863	CCCATGTGTGCTGCCCCCTCAGCTCTCTCATCTGCAAGTACGTCCACAAACACACAAC	922
QY	361	ACTTACACCAACATGAGAACTGCTCTCTGCGCTCTGCGGACAGCATGAGGAGCCGC	420
DB	923	ACTTACACCAACATGAGAACTGCTCTCTGCGCTCTGCGGACAGCATGAGGAGCCGC	982
QY	421	ACCTTTGCGCTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT	480
DB	983	ACCTTTGCGCTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT	1042
QY	481	AATGAATACAAACGGCTCTTACGTGCGACCCGCGTGAAGAGTGGGTGCGACTCTCTTAA	540
DB	1043	AATGAATACAAACGGCTCTTACGTGCGACCCGCGTGAAGAGTGGGTGCGACTCTCTTAA	1102
QY	541	AATCTCCGCTTTTATAAATACTACACGCTGTGCGAAGCGGGTGAAGAAAGACGCTCC	600
DB	1103	AATCTCCGCTTTTATAAATACTACACGCTGTGCGAAGCGGGTGAAGAAAGACGCTCC	1162
QY	601	GACTACTCAAGGATTTACCTCAGACCTCATCAATGACAGCGTGTGCTTTCCGC	660
DB	1163	GACTACTCAAGGATTTACCTCAGACCTCATCAATGACAGCGTGTGCTTTCCGC	1222
QY	661	ACGTCCAAGAGATGTACCCGCGACAGCGCAGTCTCATGCTCATGACCATGCGACGCC	720
DB	1223	ACGTCCAAGAGATGTACCCGCGACAGCGCAGTCTCATGCTCATGACCATGCGACGCC	1282
QY	721	CACGGCCCTGAGGATTCAGCCCCCAATAATTACGCTCTTCCCAAGCATCTCAGCAC	780
DB	1283	CACGGCCCTGAGGATTCAGCCCCCAATAATTACGCTCTTCCCAAGCATCTCAGCAC	1342
QY	781	ATCAGCGCGAGCTACAACTACGCGCCCAACCGGACAAACACTGGATCATGGCTACAG	840
DB	1343	ATCAGCGCGAGCTACAACTACGCGCCCAACCGGACAAACACTGGATCATGGCTACAG	1402
QY	841	GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATCTCCAGCGGAGCGCTTGCA	900
DB	1403	GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATCTCCAGCGGAGCGCTTGCA	1462
QY	901	ACCTCATGTGCGTGAAGCTTCCATGAGAGAGATTTACAACTGCTGGTTGAGACGGGC	960
DB	1463	ACCTCATGTGCGTGAAGCTTCCATGAGAGAGATTTACAACTGCTGGTTGAGACGGGC	1522
QY	961	GAGCTGACAAACACGTPACATGATATACCGCGGACCGGTTTACACATCGGCCAGTTT	1020
DB	1523	GAGCTGACAAACACGTPACATGATATACCGCGGACCGGTTTACACATCGGCCAGTTT	1582
QY	1021	GGCCTGTGAAAGGGAAATCCATGCGCATATGATTTGACATCAGGGTCCCGTTTACGTG	1080
DB	1583	GGCCTGTGAAAGGGAAATCCATGCGCATATGATTTGACATCAGGGTCCCGTTTACGTG	1642
QY	1081	AGGGGCCCCAAAGTGAAGCGGCTGTCTGATCCCGCATCTCTCAACATTCAGCTG	1140
DB	1643	AGGGGCCCCAAAGTGAAGCGGCTGTCTGATCCCGCATCTCTCAACATTCAGCTG	1702
QY	1141	GCSCCCACCATCTTGACATTTGAGCCCTTGAGCATACCTTGGGGATATGAGCGGAAATCC	1200
DB	1703	GCSCCCACCATCTTGACATTTGAGCCCTTGAGCATACCTTGGGGATATGAGCGGAAATCC	1762

QY	1201	ATCTCAAGCTGCTGACACACGAGCGCGCGTGAATCGGTTTCACCTTGA AAAAAGAAAGATG	1260
Db	1763	ATCTCAAGCTGCTGACACACGAGCGCGCGTGAATCGGTTTCACCTTGA AAAAAGAAAGATG	1822
QY	1261	AGGTCCTGGCGGACCTCTTCTTGGTGGAGAGGCAAGCTGCTACACAAGAGAGACAAT	1320
Db	1823	AGGGTCTGGCGGACCTCTTCTTGGTGGAGAGGCAAGCTGCTACACAAGAGAGACAAT	1882
QY	1321	GACAAGGTGGA CGCCACAGAGAGAGAACTTTCTGCCCAAGTACCAAGCGTGTGAAGGACCTTG	1380
Db	1883	GACAAGGTGGA CGCCACAGAGAGAGAACTTTCTGCCCAAGTACCAAGCGTGTGAAGGACCTTG	1942
QY	1381	TGTCAGCGTGTGAGTACACAGACGGCGTGTGACAGCTGGGACAGAAAGTGCGAGCTGTG	1440
Db	1943	TGTCAGCGTGTGAGTACACAGACGGCGTGTGACAGCTGGGACAGAAAGTGCGAGCTGTG	2002
QY	1441	GAGGACGCCACGGGGAGCTGAAGCTGCATAAGTGCAAGGGGCCCTATGCGGCTGGCGCG	1500
Db	2003	GAGGACGCCACGGGGAGCTGAAGCTGCATAAGTGCAAGGGGCCCTATGCGGCTGGCGCG	2062
QY	1501	ACGAGAGCCCTCTCCAAACCTCGTGCCCAAGTACTACGGCGAGGGCAGCGAGGCTGCACC	1560
Db	2063	ACGAGAGCCCTCTCCAAACCTCGTGCCCAAGTACTACGGCGAGGGCAGCGAGGCTGCACC	2122
QY	1561	TGTGACAGCGGGACTACAAGCTCAGCTGGCCGGGACGCGGGA AAAAATCTTTCAGAAG	1620
Db	2123	TGTGACAGCGGGACTACAAGCTCAGCTGGCCGGGACGCGGGA AAAAATCTTTCAGAAG	2182
QY	1621	AAGTACAAGGCCAGCTATGTCCAGTTCGGTCAATCGCTCAGTGGCCATCGAGGTGGAC	1680
Db	2183	AAGTACAAGGCCAGCTATGTCCAGTTCGGTCAATCGCTCAGTGGCCATCGAGGTGGAC	2242
QY	1681	GGCAGGGTGTACCAAGCTGAGGCTTGGGTGATGCGCGCCAGGCCGAAAACCTCACCAGCGG	1740
Db	2243	GGCAGGGTGTACCAAGCTGAGGCTTGGGTGATGCGCGCCAGGCCGAAAACCTCACCAGCGG	2302
QY	1741	CACCTGCGCAGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACATTCAGTGGCACT	1800
Db	2303	CACCTGCGCAGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGACATTCAGTGGCACT	2362
QY	1801	GGAGGCTTCCGACACTACAGCGGCCACCCCATTTAAAGTGTACACATCGGTGCTACATC	1860
Db	2363	GGAGGCTTCCGACACTACAGCGGCCACCCCATTTAAAGTGTACACATCGGTGCTACATC	2422
QY	1861	CTAGAGAACACACAGCTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCGAGGCTGGAAA	1920
Db	2423	CTAGAGAACACACAGCTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCGAGGCTGGAAA	2482
QY	1921	GACCAAGCTGCACATCGACACAGAGATTGAAACCCCTGCAGAACAAAATTAGAAGCCTTG	1980
Db	2483	GACCAAGCTGCACATCGACACAGAGATTGAAACCCCTGCAGAACAAAATTAGAAGCCTTG	2542
QY	1981	AGGGAAGTCGAGGTCACTTGAAGAAAAGCGCCAGAGAAATGTGACTGTCACAAAATC	2040
Db	2543	AGGGAAGTCGAGGTCACTTGAAGAAAAGCGCCAGAGAAATGTGACTGTCACAAAATC	2602
QY	2041	AGCTACACACCGACACAAAGCGCGCTCAAGGACAGAGGCTCCAGTGTGCAATCCTTTC	2100
Db	2603	AGCTACACACCGACACAAAGCGCGCTCAAGGACAGAGGCTCCAGTGTGCAATCCTTTC	2662
QY	2101	AGGAAGGCTGCAAGAGAAAGGACAAGTGTGCTGTTGCGGACAGAGAGCCCAAGAG	2160
Db	2663	AGGAAGGCTGCAAGAGAAAGGACAAGTGTGCTGTTGCGGACAGAGAGCCCAAGAG	2722
QY	2161	AAACTCCGCAAGCTGTCAAGCGCCTGCAGAACAAACGACACGTGCGACATGCCAGGCCTC	2220
Db	2723	AAACTCCGCAAGCTGTCAAGCGCCTGCAGAACAAACGACACGTGCGACATGCCAGGCCTC	2782
QY	2221	ACGTGCTTACCACGACACCGACACTGGCAGACGGCGCTTTCTTGACACATCGGGGCT	2280
Db	2783	ACGTGCTTACCACGACACCGACACTGGCAGACGGCGCTTTCTTGACACATCGGGGCT	2842
QY	2281	TTCTGTGCCCTGCACCGAGGCCAACAAATAACACGTACTGTGGTGCATGAGGACCAATCAATGAG	2340

2843 TTCTGTGCTGCAACGCGCCAAACAATAACACGTACTGTGTGATGAGGACCACTCATATGAG 2902

2341 ACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTCTCTAGAGTACTTTGATCTCAAC 2400

2903 ACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTCTCTAGAGTACTTTGATCTCAAC 2962

2401 ACAGACCCCTACAGCTGATGAATGACAGTGAACAACATGCGACGGGATGCTCTCAACGAG 2460

2963 ACAGACCCCTACAGCTGATGAATGACAGTGAACAACATGCGACGGGATGCTCTCAACGAG 3022

2461 CTACACGTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAACAGCTGTAAACCCCGG 2520

3023 CTACACGTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAACAGCTGTAAACCCCGG 3082

2521 ACTCGAAACATGAGCTGGGACCTTAAAGATGAGGAGAGCTATGAGCAATACAGGACGTTT 2580

3083 ACTCGAAACATGAGCTGGGACCTT-----GATGAGGAGAGCTATGAGCAATACAGGACGTTT 3133

2581 CAGCGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGACAACTGTGG 2640

3134 CAGCGCTCGAAGTGGCCAGAAATGAAGAGACCTTTCTTCCAAATCACTGGGACAACTGTGG 3193

2641 GAAGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCTCAAAAACATAGAGGATCAAC 2700

3194 GAAGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCTCAAAAACATAGAGGATCAAC 3253

2701 TGACTGCACAGGCAATGAAAAACATGTGGGTGATTTCCAGCAGACCTGTGCTATTTGGCC 2760

3254 TGACTGCACAGGCAATGAAAAACATGTGGGTGATTTCCAGCAGACCTGTGCTATTTGGCC 3313

2761 AGGAGGCTGAGAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATTAAC 2820

3314 AGGAGGCTGAGAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATTAAC 3373

2821 CAGCAGGACGAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTA 2880

3374 CAGCAGGACGAGATTAACCTTCAGGAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTA 3433

2881 TACCTCACCAGCTGCACAAAATGCATTTTGTATCAAAAAGTCCACATPAAACCTCTCC 2940

3434 TACCTCACCAGCTGCACAAAATGCATTTTGTATCAAAAAGTCCACATPAAACCTCTCC 3493

2941 CAGAGAGCTCAAAAGAAAACGAGAGAGCGAGCGAGAGATTTCTCTGGAAATTTCT 3000

3494 CAGAGAGCTCAAAAGAAAACGAGAGAGCGAGCGAGAGATTTCTCTGGAAATTTCT 3553

3001 CCCAAGGCGAAAAGTCAATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTCTTAAA 3060

3554 CCCAAGGCGAAAAGTCAATGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTCTTAAA 3613

3061 TGTCTTATTTCTTTTGTGTCAAAAGAGGAATCAAGAGCAGGACAGAGGACAGT 3120

3614 TCTCTTATTTCTTTTGTGTGTCAAAAGAGGAATCAAGAGCAGGAACTAGAGCAGGACAGT 3673

3121 CGAGAGGCTGAAAACAGTGCAGAGAGCTTTGACAATGAGTCAAGTACGACAAAAGAGATGA 3180

3674 GAGAGGCTGAAAACAGTGCAGAGAGCTTTGACAATGAGTCAAGTACGACAAAAGAGATGA 3733

3181 CATTTACTAGCACTATAAACCTGTGTGCTCTGAGAGAACTGCCTTCATTGTATATAT 3240

3734 CATTTACTAGCACTATAAACCTGTGTGCTCTGAGAGAACTGCCTTCATTGTATATAT 3793

3241 GTGACATTTTATGATTAATCAACATGGAACCTTTTAGGGGAACCTTAATAAGAAATCCCAA 3300

3794 GTGACATTTTATGATTAATCAACATGGAACCTTTTAGGGGAACCTTAATAAGAAATCCCAA 3853

3301 TTTTTCAGAGTGGTGGTGTCAATAAACCGTCTGTGGCCAGTGTAAAAAGAAAA 3353

3854 TTTTTCAGAGTGGTGGTGTCAATAAACCGTCTGTGGCCAGTGTAAAAAGAAAA 3906

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 14:35:48 ; Search time 221 Seconds
(without alignments)
6716.613 Million cell updates/sec

Title: US-10-025-966A-4

Perfect score: 3363

Sequence: 1 gactccgcgctcccaaga.....aaagaaaaaaaaaa 3363

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued Patents NA.*
- 2: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	3323.4	98.8	3906	4	US-09-996-243-83	Sequence 83, Appl
2	1308.6	38.9	3057	4	US-09-886-319A-53	Sequence 53, Appl
3	1073	31.9	5769	4	US-09-668-673B-1	Sequence 1, Appl
4	1040.4	30.9	2616	4	US-09-773-426A-12	Sequence 12, Appl
5	1040.4	30.9	4321	4	US-09-773-426A-4	Sequence 4, Appl
6	1002.4	29.8	4834	4	US-09-668-673B-15	Sequence 15, Appl
7	1001.6	29.8	1367	4	US-09-668-673B-13	Sequence 13, Appl
8	578.2	17.2	1799	4	US-09-810-347-1	Sequence 1, Appl
9	441	13.1	441	4	US-09-404-879A-315	Sequence 315, Appl
10	381.6	11.3	494	4	US-09-668-673B-10	Sequence 10, Appl
11	375.8	11.2	466	4	US-09-668-673B-9	Sequence 9, Appl
12	369	11.0	3871	4	US-09-484-970B-70	Sequence 70, Appl
13	368	10.9	459	4	US-09-668-673B-12	Sequence 12, Appl
14	347.4	10.3	538	4	US-09-668-673B-8	Sequence 8, Appl
15	297.4	8.8	436	4	US-09-668-673B-11	Sequence 11, Appl
16	291	8.7	296	4	US-09-668-673B-21	Sequence 21, Appl
17	288.4	8.6	302	4	US-09-313-294A-7135	Sequence 7135, Appl
18	210.8	6.3	540	4	US-09-668-673B-5	Sequence 5, Appl
19	177	5.3	482	4	US-09-668-673B-6	Sequence 6, Appl
20	172	5.1	3577	4	US-09-620-312D-457	Sequence 457, Appl
21	157.4	4.7	288	4	US-09-668-673B-19	Sequence 19, Appl
22	156.6	4.7	42571	4	US-09-810-347-3	Sequence 3, Appl
23	104	3.1	590	4	US-09-668-673B-17	Sequence 17, Appl
24	99.4	3.0	614	4	US-09-668-673B-1	Sequence 1, Appl
25	74	2.2	473	4	US-09-668-673B-4	Sequence 4, Appl
26	59.2	1.8	7218	1	US-08-232-463-14	Sequence 14, Appl
27	50.4	1.5	1859	3	US-08-894-818B-15	Sequence 15, Appl

28	50.4	1.5	1977	3	US-08-894-818B-2	Sequence 2, Appl
29	50.4	1.5	1977	3	US-08-894-818B-6	Sequence 6, Appl
30	50.4	1.5	1977	4	US-09-445-472-11	Sequence 11, Appl
31	49.4	1.5	505	4	US-08-961-527-389	Sequence 389, Appl
32	49.4	1.5	5552	3	US-08-153-888-1	Sequence 1, Appl
33	48.8	1.5	290	4	US-09-444-791A-113	Sequence 113, Appl
34	48.8	1.5	1926	4	US-09-249-585A-2	Sequence 2, Appl
35	48.8	1.5	1926	4	US-09-410-399-3	Sequence 3, Appl
36	48.8	1.5	2580	3	US-09-050-863-2	Sequence 2, Appl
37	48.8	1.5	2580	4	US-09-359-081-2	Sequence 2, Appl
38	48.8	1.5	5452	2	US-09-130-114-1	Sequence 1, Appl
39	48.8	1.5	8705	4	US-09-647-344A-14	Sequence 14, Appl
40	48.8	1.5	9600	3	US-08-910-647-1	Sequence 1, Appl
41	48.8	1.5	9600	4	US-09-620-925-1	Sequence 1, Appl
42	48.8	1.5	10596	1	US-07-884-811-15	Sequence 15, Appl
43	48.8	1.5	10596	1	US-07-885-971-15	Sequence 15, Appl
44	48.8	1.5	10596	1	US-08-087-783A-15	Sequence 15, Appl
45	48.8	1.5	10596	1	US-08-194-088B-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-996-243-83
; Sequence 83, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28

1	PRIOR APPLICATION NUMBER: 60/084560
2	PRIOR FILING DATE: 1998-05-07
3	PRIOR APPLICATION NUMBER: 60/087106
4	PRIOR FILING DATE: 1998-05-28
5	PRIOR APPLICATION NUMBER: 60/087607
6	PRIOR FILING DATE: 1998-06-02
7	PRIOR APPLICATION NUMBER: 60/087609
8	PRIOR FILING DATE: 1998-06-02
9	PRIOR APPLICATION NUMBER: 60/087759
10	PRIOR FILING DATE: 1998-06-02
11	PRIOR APPLICATION NUMBER: 60/087827
12	PRIOR FILING DATE: 1998-06-03
13	PRIOR APPLICATION NUMBER: 60/088021
14	PRIOR FILING DATE: 1998-06-04
15	PRIOR APPLICATION NUMBER: 60/088025
16	PRIOR FILING DATE: 1998-06-04
17	PRIOR APPLICATION NUMBER: 60/088026
18	PRIOR FILING DATE: 1998-06-04
19	PRIOR APPLICATION NUMBER: 60/088028
20	PRIOR FILING DATE: 1998-06-04
21	PRIOR APPLICATION NUMBER: 60/088029
22	PRIOR FILING DATE: 1998-06-04
23	PRIOR APPLICATION NUMBER: 60/088030
24	PRIOR FILING DATE: 1998-06-04
25	PRIOR APPLICATION NUMBER: 60/088033
26	PRIOR FILING DATE: 1998-06-04
27	PRIOR APPLICATION NUMBER: 60/088326
28	PRIOR FILING DATE: 1998-06-04
29	PRIOR APPLICATION NUMBER: 60/088167
30	PRIOR FILING DATE: 1998-06-05
31	PRIOR APPLICATION NUMBER: 60/088202
32	PRIOR FILING DATE: 1998-06-05
33	PRIOR APPLICATION NUMBER: 60/088212
34	PRIOR FILING DATE: 1998-06-05
35	PRIOR APPLICATION NUMBER: 60/088217
36	PRIOR FILING DATE: 1998-06-05
37	PRIOR APPLICATION NUMBER: 60/088655
38	PRIOR FILING DATE: 1998-06-09
39	PRIOR APPLICATION NUMBER: 60/088734
40	PRIOR FILING DATE: 1998-06-10
41	PRIOR APPLICATION NUMBER: 60/088738
42	PRIOR FILING DATE: 1998-06-10
43	PRIOR APPLICATION NUMBER: 60/088742
44	PRIOR FILING DATE: 1998-06-10
45	PRIOR APPLICATION NUMBER: 60/088810
46	PRIOR FILING DATE: 1998-06-10
47	PRIOR APPLICATION NUMBER: 60/088824
48	PRIOR FILING DATE: 1998-06-10
49	PRIOR APPLICATION NUMBER: 60/088826
50	PRIOR FILING DATE: 1998-06-10
51	PRIOR APPLICATION NUMBER: 60/088858
52	PRIOR FILING DATE: 1998-06-11
53	PRIOR APPLICATION NUMBER: 60/088861
54	PRIOR FILING DATE: 1998-06-11
55	PRIOR APPLICATION NUMBER: 60/088876
56	PRIOR FILING DATE: 1998-06-11
57	PRIOR APPLICATION NUMBER: 60/089105
58	PRIOR FILING DATE: 1998-06-12
59	PRIOR APPLICATION NUMBER: 60/089440
60	PRIOR FILING DATE: 1998-06-15
61	PRIOR APPLICATION NUMBER: 60/089512
62	PRIOR FILING DATE: 1998-06-16
63	PRIOR APPLICATION NUMBER: 60/089514
64	PRIOR FILING DATE: 1998-06-16
65	PRIOR APPLICATION NUMBER: 60/089532
66	PRIOR FILING DATE: 1998-06-17
67	PRIOR APPLICATION NUMBER: 60/089538
68	PRIOR FILING DATE: 1998-06-17
69	PRIOR APPLICATION NUMBER: 60/089598
70	PRIOR FILING DATE: 1998-06-17
71	PRIOR APPLICATION NUMBER: 60/089599
72	PRIOR FILING DATE: 1998-06-17
73	PRIOR APPLICATION NUMBER: 60/089600

1	PRIOR FILING DATE: 1998-06-17
2	PRIOR APPLICATION NUMBER: 60/089653
3	PRIOR FILING DATE: 1998-06-17
4	PRIOR APPLICATION NUMBER: 60/089801
5	PRIOR FILING DATE: 1998-06-18
6	PRIOR APPLICATION NUMBER: 60/089907
7	PRIOR FILING DATE: 1998-06-18
8	PRIOR APPLICATION NUMBER: 60/089908
9	PRIOR FILING DATE: 1998-06-18
10	PRIOR APPLICATION NUMBER: 60/089947
11	PRIOR FILING DATE: 1998-06-19
12	PRIOR APPLICATION NUMBER: 60/089948
13	PRIOR FILING DATE: 1998-06-19
14	PRIOR APPLICATION NUMBER: 60/089952
15	PRIOR FILING DATE: 1998-06-19
16	PRIOR APPLICATION NUMBER: 60/090246
17	PRIOR FILING DATE: 1998-06-22
18	PRIOR APPLICATION NUMBER: 60/090252
19	PRIOR FILING DATE: 1998-06-22
20	PRIOR APPLICATION NUMBER: 60/090254
21	PRIOR FILING DATE: 1998-06-22
22	PRIOR APPLICATION NUMBER: 60/090349
23	PRIOR FILING DATE: 1998-06-23
24	PRIOR APPLICATION NUMBER: 60/090355
25	PRIOR FILING DATE: 1998-06-23
26	PRIOR APPLICATION NUMBER: 60/090429
27	PRIOR FILING DATE: 1998-06-24
28	PRIOR APPLICATION NUMBER: 60/090431
29	PRIOR FILING DATE: 1998-06-24
30	PRIOR APPLICATION NUMBER: 60/090435
31	PRIOR FILING DATE: 1998-06-24
32	PRIOR APPLICATION NUMBER: 60/090444
33	PRIOR FILING DATE: 1998-06-24
34	PRIOR APPLICATION NUMBER: 60/090445
35	PRIOR FILING DATE: 1998-06-24
36	PRIOR APPLICATION NUMBER: 60/090472
37	PRIOR FILING DATE: 1998-06-24
38	PRIOR APPLICATION NUMBER: 60/090535
39	PRIOR FILING DATE: 1998-06-24
40	PRIOR APPLICATION NUMBER: 60/090540
41	PRIOR FILING DATE: 1998-06-24
42	PRIOR APPLICATION NUMBER: 60/090542
43	PRIOR FILING DATE: 1998-06-24
44	PRIOR APPLICATION NUMBER: 60/090557
45	PRIOR FILING DATE: 1998-06-24
46	PRIOR APPLICATION NUMBER: 60/090676
47	PRIOR FILING DATE: 1998-06-25
48	PRIOR APPLICATION NUMBER: 60/090695
49	PRIOR FILING DATE: 1998-06-25
50	PRIOR APPLICATION NUMBER: 60/090696
51	PRIOR FILING DATE: 1998-06-25
52	PRIOR APPLICATION NUMBER: 60/090690
53	PRIOR FILING DATE: 1998-06-25
54	PRIOR APPLICATION NUMBER: 60/090694
55	PRIOR FILING DATE: 1998-06-25
56	PRIOR APPLICATION NUMBER: 60/090695
57	PRIOR FILING DATE: 1998-06-25
58	PRIOR APPLICATION NUMBER: 60/090696
59	PRIOR FILING DATE: 1998-06-25
60	PRIOR APPLICATION NUMBER: 60/090862
61	PRIOR FILING DATE: 1998-06-26
62	PRIOR APPLICATION NUMBER: 60/090863
63	PRIOR FILING DATE: 1998-06-26
64	PRIOR APPLICATION NUMBER: 60/091360
65	PRIOR FILING DATE: 1998-07-01
66	PRIOR APPLICATION NUMBER: 60/091478
67	PRIOR FILING DATE: 1998-07-02
68	PRIOR APPLICATION NUMBER: 60/091544
69	PRIOR FILING DATE: 1998-07-01
70	PRIOR APPLICATION NUMBER: 60/091519
71	PRIOR FILING DATE: 1998-07-02
72	PRIOR APPLICATION NUMBER: 60/091626
73	PRIOR FILING DATE: 1998-07-02
74	PRIOR APPLICATION NUMBER: 60/091633
75	PRIOR FILING DATE: 1998-07-02

QY 2041 AGTACACACACCCAGCACAAGGCGCCCTCAAGCAGACAGAGGCTCCAGTCTGCATCTCTTC 2100
Db 2603 AGCTACACACACCCAGCACAAGGCGCCCTCAAGCAGACAGAGGCTCCAGTCTGCATCTCTTC 2662
QY 2101 AGGAAGGCGCTGCAAGAGAAGGACAAGGTGTGGCTGTTGGGAGCAGAGCGAAGAAG 2160
Db 2663 AGGAAGGCGCTGCAAGAGAAGGACAAGGTGTGGCTGTTGGGAGCAGAGCGAAGAAG 2722
QY 2161 AAATCCCGAGCTGCTCAAGCGCTGCAAGACAGACAGACAGCAGTGCAGCAGTGCAGGCTC 2220
Db 2723 AAATCCCGAGCTGCTCAAGCGCTGCAAGACAGACAGACAGCAGTGCAGCAGTGCAGGCTC 2782
QY 2221 AGTGCCTTCACCCAGCAGCAACAGCAGCTGCGAGCGCGCTTCTTGACACTGGGCGCT 2280
Db 2783 AGTGCCTTCACCCAGCAGCAACAGCAGCTGCGAGCGCGCTTCTTGACACTGGGCGCT 2842
QY 2281 TTCTGTGCTGACAGCGCCCAACATACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2843 TTCTGTGCTGACAGCGCCCAACATACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2902
QY 2341 ACTCAAAATTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2903 ACTCAAAATTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962
QY 2401 ACAGACCCCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2963 ACAGACCCCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3022
QY 2461 CTACAGCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 3023 CTACAGCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3082
QY 2521 ACTCGAAACATGAGCCTGGGACTTAAAGATGAGGAGGATGATGAGCAATACAGGAGTTT 2580
Db 3083 ACTCGAAACATGAGCCTGGGACTTAAAGATGAGGAGGATGATGAGCAATACAGGAGTTT 3133
QY 2581 CAGCTGCAAGTGGCCAGAAATGAGGACCTTCTTCCAAATCACTGGGACACTGTTG 2640
Db 3134 CAGCTGCAAGTGGCCAGAAATGAGGACCTTCTTCCAAATCACTGGGACACTGTTG 3193
QY 2641 GAAGGCTGGGAAGTTTAAAGAAACACAGAGGTGGACCTCCAAAAACATAGAGGATCAC 2700
Db 3194 GAAGGCTGGGAAGTTTAAAGAAACACAGAGGTGGACCTCCAAAAACATAGAGGATCAC 3253
QY 2701 TGACTGCAGAGCAATGAAAACCATGTTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3254 TGACTGCAGAGCAATGAAAACCATGTTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313
QY 2761 AGGAGGCTGAGAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3314 AGGAGGCTGAGAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
QY 2821 CAGCAGAGCAGAGATPAACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTA 2880
Db 3374 CAGCAGAGCAGAGATPAACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTTGGATTA 3433
QY 2881 TACCTCAGCAGCTGCACAAATGCAATTTTTCGTATCAAAAAGTCCACCACTAACCTCCC 2940
Db 3434 TACCTCAGCAGCTGCACAAATGCAATTTTTCGTATCAAAAAGTCCACCACTAACCTCCC 3493
QY 2941 CCAGAAGCTCACAAGGAAACCGAGAGAGCGAGAGAGATTTCCTTTGGAAATTTCT 3000
Db 3494 CCAGAAGCTCACAAGGAAACCGAGAGAGCGAGAGAGATTTCCTTTGGAAATTTCT 3553
QY 3001 CCAAGGCGGAAGTCAATGGATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAA 3060
Db 3554 CCAAGGCGGAAGTCAATGGATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAA 3613
QY 3061 TCCTCTTATTTCTTTGTTGTGTCACAAAGAGCACTAAAGAGCAGGACAGAGGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTGTTGTGTCACAAAGAGCACTAAAGAGCAGGACAGAGGCAACGT 3673

QY 3121 CGAGAGGCTGAAAAACAGTGCAGAGACGTTTGAACAATGAGTCAGTACACAAAAAGATGA 3180
Db 3674 CGAGAGGCTGAAAAACAGTGCAGAGACGTTTGAACAATGAGTCAGTACACAAAAAGATGA 3733
QY 3181 CATTTACCTAGCACTATAAAACCTCGTTCCTCTCTGAGAAACTGCTTCATTTGTTATAT 3240
Db 3734 CATTTACCTAGCACTATAAAACCTCGTTCCTCTCTGAGAAACTGCTTCATTTGTTATAT 3793
QY 3241 GTGACTATTATCATGTAAATCAACATGGGAACCTTTTAGGGAACTTAAAGAAATCCCAA 3300
Db 3794 GTGACTATTATCATGTAAATCAACATGGGAACCTTTTAGGGAACTTAAAGAAATCCCAA 3853
QY 3301 TTTTCAGGAGTGGTGTGTCATAAAACGCTCTGTGCGCCAGTGTAAAAAGAAAA 3353
Db 3854 TTTTCAGGAGTGGTGTGTCATAAAACGCTCTGTGCGCCAGTGTAAAAAGAAAA 3906

RESULT 2
US-09-886-319A-53
; Sequence 53, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regendogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the diagnosis or treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 3057
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-886-319A-53

Query Match 38.9%; Score 1308.6; DB 4; Length 3057;
Best Local Similarity 86.0%; Pred. No. 0;
Matches 1549; Conservative 1; Mismatches 225; Indels 26; Gaps 8;

QY 943 ATGCTGTTGAGACGGGCGAGCTGGACAACAGTACATCGTATACACCGCCGACCAACGGT 1002
Db 2 ATGCTGTTGAGACGGGCGAGCTGGACAACAGTACATCGTATACACCGCCGACCAACGGC 61
QY 1003 TACCAATGCGCCAGTTTGGCTGTGTTGAAAGGAAATCATGTCATATGAGTTTGATC 1062
Db 62 TACCAATGCGCCAGTTTGGCTGTGTTGAAAGGCAAGTCTATGCGGTATGATTCGACATC 121
QY 1063 AGGTCCTGCTTACCTGAGGGGCCCCAACGTGGAAGCGGCTGTCTGAATCCCCACATC 1122
Db 122 AGAGTCCCGTTTACCTGAGGGGCCCCAACGTGGAAGCTGGCTCTCTGAACCCCCACAT 181
QY 1123 GTCTCAACATTGACCTTGGCCCCCAACCTTCTGGAATTTGAGGCTTGGACATCTGCG 1181
Db 182 GTCTCAACATTGACCTTGGCCCCCAACCTTCTGGAATTTGAGGCTTGGACATCTGCG 241
QY 1182 GGATATGAGCGGGAATTCATCTCAAGCTGCTGGAACCGGCGCTGCTGAATCCGTT 1241
Db 242 AGCATGAGCGGGAAGTCTTATCTCAAACTTCTGGAATTTGAGGCTTGGACCGGTT 301
QY 1242 TCATTTGAAAAAGAGATGAGGGTCTGCGGAGCTCTTTCTTTGGTGGAGAGGCAAGC 1300
Db 302 CCATTTGAAAAAGAGTGAAGGCTTGGCGAGCTCTTCTCTGTTGGAGAGGCAAGC 361

1301 -TGCTACACAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTTCTGCCCAAG 1359
1362 TTGCTCCACACAGAGAGAGGAGTGAAGTGAATGCCAGAGGAGAACTTTCTGCCCAAG 421
1360 TACACAGGTGTGAAGAACTGTGTGAGGTGTGAGTACAGAGCGGTGTGAGAGAGCTG 1419
422 TACACAGCGGTGAAGAACTGTGTGAGGTGTGAGTACAGAGCGGTGTGAGAGAGCTG 481
1420 GACAGAGAGTGCAGTGTGTGAGAGAGCGGAGGAGCTGAGAGTGCATATAGTGAAG 1479
482 GGGCAGAGTGCAGTGTGTGAGAGAGCGTTCGAGAGCGCTGAGAGTGCATATAGTGA 541
1480 GGGCCCATCGGCT-----GGGCGGACAGAGAGCGCTCTCCAACTGTGTGCCCAAG 1530
542 GGGCCCATCGGCTTTGGTGGCGGCGGTGGCAGAGAGCGCTCTCCAACTGTGTGCCCAAG 601
1531 TACTAGGGGAGGAGAGCGGCTGCACTGTGACA-----GGGCGGAGTACAG 1581
602 TATGACGGCCAGAGAGAGCGGCTGCACTGTGACA-----GGGCGGAGTACAG 661
1582 CTCAGCGCTGGCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1641
662 CTGGGCGCTGGTGGAGCGG---TAACTCTTTAAGAAAGTATAGACCAAGCTATGCC 718
1642 CGCAGTGTGCTCATCGCTCAGTGGCCATGAGGTGGAGCGGAGTGTATCCAGTAGGC 1701
719 CGGAACCGCTCATCGCTTCGTTGGTGGCCATCGAGGTGGAGCGGAGTGTATCCAGTAGGC 778
1702 CTGGGTGTATGCGCGCCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1761
779 TTGGATGTGTGCTCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
1762 GACCAAGATGACAAGATGGTGGGAGTTCAGTGGCACTGAGGAGGAGGAGGAGGAGGAGGAG 1821
839 GACCAAGATGACAAGATGGTGGGAGTTCAGTGGCACTGAGGAGGAGGAGGAGGAGGAGGAG 898
1822 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1881
899 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
1882 TGTGACCTGGAGCTGTACAAGTCCCTGAGGCGCTGGAAGAGGAGGAGGAGGAGGAGGAGGAG 1941
959 TCGAGTGTGAGCTGTACAAGTCCCTGAGGCGCTGGAAGAGGAGGAGGAGGAGGAGGAGGAG 1018
1942 CAGGAGTGTGAACCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2001
1019 CATGAGATCGAAACCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078
2002 AAGAAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2061
1079 AAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1138
2062 GCGGCGCTCAAGCAGAGAGGCTTCCAGTGTGATCTGATCTGATCTGATCTGATCTGATCTGAT 2121
1139 GCGGCGCTCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198
2122 GACAGGTGTGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2181
1199 GACAGGTGTGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1256
2182 GCGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2241
1257 GCGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1316
2242 CAGGAGTGTGCGGCTTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2301
1317 CAGGAGTGTGCGGCTTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1376
2302 AACAAATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2361
1377 AACAAATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1436

2362 GAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAACACAGAGAGGAGGAGGAGGAGGAGGAG 2421
1437 GAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAACACAGAGAGGAGGAGGAGGAGGAGGAG 1496
2422 ATGAGTGTGAACACACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2481
1497 AAGCGGTGAACACACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1556
2482 CTGAGGAGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2541
1557 CTAAGGAGCTGTAAAGGCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616
2542 CTTAAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2601
1617 CTTAGAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1676
2602 ATGAGAGAGCTTCTTCCAAATCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2661
1677 ATGAGAGAGCTTCTTCCAAATCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1736
2662 ACAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2721
1737 GCCATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1796
2722 A 2722
1797 A 1797

RESULT 3
US-09-668-673B-1
; Sequence 1, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5769
; TYPE: DNA
; ORGANISM: Quails
US-09-668-673B-1

Query Match 31.9%; Score 1073; DB 4; Length 5769;
Best Local Similarity 65.9%; Pred. No. 1.3e-265;
Matches 1664; Conservative 4; Mismatches 809; Indels 48; Gaps 6;

97 CTGCTGGGTTGAAGCTCGGCTTCTGTCGACACACCGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAG 156
788 CTGCTGACAGTCAATCTTCCACTCTCAAGTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
157 GACCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 216
848 GAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 907
217 CTGGGTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 276
908 CTAGGAGTCTTCAAGTGAATGAACAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 967
277 TTATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 336
968 TTATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027
337 GGCAGTACGTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 396

Db 1028 GGAAGATGATGCAACCAACACATATACACCAACATGAACCTCTCTCTCCCTCC 1087
QY 397 TGGGAGGACAGCAGAGAGCGGACCTTTGGCGGTGTACCTCAATAGACGTGGGTACCGG 456
Db 1088 TGGGAGGCACTCAGAGGACGCGACCTTTGGCGGTGTATCTGAATAACACTGGGTATCGA 1147
QY 457 ACAGCTTTCTTGGGAAGTATCTTAATCAATACACAGGCTCTCTAGTGCACCCGGCTGG 516
Db 1148 ACAGCTTTTGGGAAGTATCTTAATCAATACACAGGCTCTCTAGTGCACCCGGCTGG 1207
QY 517 AAGGAGTGGTGGGACTCTTTAAATAACCTCCGCTTTTATTAACACAGCTGTGTGGAAC 576
Db 1208 AGAGAGTGGTGGATTAAGTGAAGAACTCTCGCTCTTATTAATACACCAATTTCTCGCAAT 1267
QY 577 GGGGTGAAGAAAGACAGGCTCCGACTACTCCAGGATTAACCTACAGACCTCATCACC 636
Db 1268 GGTAAACAAAGAGACATGGATTTGATTAATGAAGGACTACTTCAACAGACTTAATCACT 1327
QY 637 AATGACAGCTGAGCTTTCTCGACCTCCAGAGATGTACCGGACACAGGCTCTCTC 696
Db 1328 AATGAGAGCAATTAATTAATCTCAGAAATGTCCAGAGGATATACCCACATAGGCCATAATG 1387
QY 697 ATGGCTCATCAGCATGACGCCCCCAGCGGCTGAGGATTCAGCCCCACATATTTACGC 756
Db 1388 ATGGCTCATCAGCATGCTCGGCTCTATGGCTCTGAGGATTCGGCCCCACATTTCTCAGAG 1447
QY 757 CTCCTTCCAAAGCGCATCTCAGCACATCAGCGGAGCTACAACTAGCGGCCCAACCGGAC 816
Db 1448 CTCCTCCCAACGCTTCAAGCATATACCCCGAGCTATACTATGACCAACATGGAT 1507
QY 817 AAACACTGGATCATGGCTTACAGGGGGCCATGAAGCCCATCCACATGAATTCACCAAC 876
Db 1508 AAGCACTGGATCATGAGTACAGGGGGCCATGTGSCCTTCCACATGGAGTTTACAAAC 1567
QY 877 ATGCTCCAGCGGAAGCGCTTGCAGACCTCTCATGTGGTGGAGCTCCATGGAGACGAT 936
Db 1568 GTCTTGCACGCAAGAGACTTCAGACCTCTGATGTGATGACTCTATGGAAGATTA 1627
QY 937 TACACATGCTGGTGAAGCGGCGAGCTGGAACAACAGTATCATGATATACCGCGGAC 996
Db 1628 TACCAGATGCTTGCAGAAATGGGAGAGCTGGAGATACCTCATATTTTACACAGCTGAC 1687
QY 997 CAGGTTACCATCGGCGAGTTGGCTGTGGAAGGGAATCCATCCCATATGAGTTT 1056
Db 1688 CATGTTACCATATGGCGAGTTTGAATGCTGTCAGGGAAGTCAATGCCATATGACTTT 1747
QY 1057 GACATCAGGCTCCGTTCTACGTGAGGGGCCCAACAGTGAAGCGGCTGTCTGAATCCC 1116
Db 1748 GATATTCGAGTTCTTCTTTATTCGTGTCAGGTAGAGCCAGGATCTGTAGTGCT 1807
QY 1117 CACATGCTCTCAACATGACCTGGCCCCCAGCTCTGGAATTCAGGCTGAGGCTGACATA 1176
Db 1808 CAGATAGTTCTGAATATGATCTTGACCAACAATCTGGAATTCAGGACTTGAACACA 1867
QY 1177 CTTGCGGATATGACGGGAAATCCATCTCTCAAGCTGCTGGAACGAGCGCGGTGAAT 1236
Db 1868 CTTCCAGATATGATGCGCAATCTGTCTTAAAGCTTCTAGACTTGGAGAGACAGGAAT 1927
QY 1237 CGGTTTCACTTGAAGAAAGATGAGGCTCTGGCGGACTCTTCTTGGTGGAGAGGCG 1296
Db 1928 AGGTTTCGAACAAACAGAGGACCAAAATCTGGCGTGACATCTCTGGTGAAGAGGCG 1987
QY 1297 AAGCTGCTACAAAGAGACATGACAGGTTGGAGCGGCCAGGAGGAGAACTTTCTGCCC 1356
Db 1988 AATTTCTCGCAAAAGAGGAGCTTACAAAACACTCAGCAATCTAATCAACTGCCA 2047
QY 1357 AAGTACAGGCTGTGAAGACCTGTGTAGCGTGTGAGTACACAGCGGCTGTGAGCAG 1416
Db 2048 AAGTATGAGAGGGTAAAGAAATTTATCCAAACAGCGAGATACACAGACGCTGTGAACA 2107
QY 1417 CTGGGACAGAGTGGCAGTGTGTGGAGGAGCGCCACGGGGAAGCTCAAGCTGATAGTGC 1476

Db 2108 CCAGGACAGAGTGGCAGTGCACAGAGATGCTTCTGGCAAGCTTCGAATTCACAAGTGC 2167
QY 1477 AAGGCCCCCA---TGGGCTGGCGGACGAGAGCCCTCTCAACCTCTGTCGCCCAAGTAC 1533
Db 2168 AAGGTATCTAGTGAATCTCTGCCATCAGAAAGGACCCGAGCATCTCACTCCAGGGA 2227
QY 1534 TACGGGACAGGAGGAGGCTCTGACCTGTGACAGCGGGACTACAAGCTCAGCCTTGCC 1593
Db 2228 TAVAGTGGTAAAGATAAGGACTGCACTGTGGAGACACCGATTTCCGAAAACAGCAGGACC 2287
QY 1594 GAGCG-----CGHAAAACCTTCAAGAAAGTCTCAAGGAGTCTCAAGGCCAGC 1635
Db 2288 CAAAGAAAATAACAAAGGAGCTTTCTGAGAAACCCAGTGGCGCAAAAATACAAACACGT 2347
QY 1636 TATGTCGCGAGTCCGCTCATCGCTCAGTGGCCATCGAGGTGGACGGCAGGGGTGTACAC 1695
Db 2348 TTTGTTTCACTCGCCAAACCCGGTCTTGTGAGTGAATTTGAAGGTGAATAATATGAC 2407
QY 1696 GTAGGCTGGTGTGCGGCC-----AGCCCCGAAACCTCACAGGCGGAC 1743
Db 2408 ATAAACCTCGAAGAGGAAGAACTGAGGTGTTAAAGACCAAGATATCACAAACGTCAC 2467
QY 1744 TGGCCAGGGGCCCTGAGGACCAAGATGACAGGATGG-----TGGGACTTCAAGTGC 1797
Db 2468 AATGCTGAAAATGACAAAAGAGGAACTGATGCTCTCTGTTGACACGATGTT 2527
QY 1798 ACTGAGGCTTCCCGACTACTCAGCGCCAAACCCATTAAGTGCACATCGGTGCTAC 1857
Db 2528 GCTGATGGCACTGATTTATAGGTCAACCCAGTCTGTGAGAGTGAACACAGTGT 2587
QY 1858 ATCTAGAGAAACGACACAGTCCAGTGTGACCTGACCTGTACAAGTCTCTGAGGCTGG 1917
Db 2588 ATTCTTCAAATGACACTTCTCTGTGAGGAGCTGTACCAATCTGCCAGACCTGG 2647
QY 1918 AAGACCAACAGCTGCATCGAACCAAGAGATTTAAACCTTCGACAAACAAATTAAGAAC 1977
Db 2648 AAGACCAACAGGCTTACATCGATAAGGAGATTTAAGCTCTCCAGGACAAAATCAAGAT 2707
QY 1978 CTGAGGAGTCCGAGGTCACTGAGAAAAGCGCCAGAGAAATGTGACTGTCAAAA 2037
Db 2708 TTGAGGAAAGTTAGAGGACACCTTAAAGAGAAACACAGACGATGTGACTACTAAA 2767
QY 2038 ATCAGTACCAACCCAGCACAAAGGC-----CGCTCAAGCAGAGGCTTCAAGTGTG 2091
Db 2768 CAGAGCTTACTACAAAGAGAAAGCGGTAAAGACCAAGAGAAAATCAAGAGCCATCTA 2827
QY 2092 CATCTTTTCAAGAGCGCTGCAAGAGAGG---CAAGGTGTGCTGTGCGGAGCAG 2148
Db 2828 CATCTCTTCAAGAGAGCAGCAGGAGGTAGACAGCAAACTGAGCTGTTCAGAGAAAT 2887
QY 2149 AAGCGCAAGAGAAACTCCGCAAGCTGCTCAAGCGCTGACAGAAACAAACGACAGTGCAGC 2208
Db 2888 CGCAGAGAGAGAGAGAAAGAGGAAAAGCGCCAGAAAGGGGATGAGTGTAGC 2947
QY 2209 ATGCCAGGCTCAGGTCTTCCACAGCAACAGAGCTGCGCAGACGGCGCTTCTCG 2268
Db 2948 CTTCTGAGCTGACATGTTTACTCATGCAATAACCACTTGGCAAACTGCACTTCTCG 3007
QY 2269 ACAGTGGGCTTTCTGTGCTGCAAGCGCAACAAATAACGCTACTGTGTCATGAGG 2328
Db 3008 AACTTGGATCTTCTGTGCTGCAAGCTCAATTAACAACTTACTGTGTTTGGCA 3067
QY 2329 ACCATCAATGAGACTCAATTTCTCTCTGTGAAATTTGAACTGGCTTCTTAGAGTAC 2388
Db 3068 ACAGTGAATGACACCCACAAATTTCTCTTGTGAAATTTGCAACTGGCTTCTTGGAAATWC 3127
QY 2389 TTTGATCTCAACACAGACCCCTTACAGCTGATGAATGCTGAGTGAACACTGGCTTCTTAGAGTAC 2448
Db 3128 TTTGATGATGACACTGACCCCTTATGAGCTGACAAATACCGTACTACATGAGTGAAGGCG 3187
QY 2449 GTCTCAACAGCTACACGCTACAGCTCATGGAGCTGAGAGCTGCAAGGCTTACAGCAG 2508
Db 3188 ATTTTAAATCAATTAATGATGATGATTAATGGAATTAACGAAGTTGTCAAGGTTTATAAGCAG 3247

1697 TAGGCTGGTGATGCG-----CCAGCCCGAAGACCTCAACAGCGCG 1741
1670 TAAATCTGGAGAGAGAGAAATTCAGAGTTTCAACCAAGAAACATTCGTAGCGTC 1729
1742 ACTGGCCAGGGGCCCTTGAGGACCAAGATGACAAAGGATGGTGGGACCTTCAGTGCGACTG 1801
1730 ATGATGAAGGCCCAAGAGGCCCAAGAGATCTCCAGGCTTCCAGTGCGTGGCAACAGGGCA 1789
1802 GAGGCTTCCCGACTACTCAGCGCC-----NACCCATTAAAGTGACACATC 1849
1790 GGATGTGGGAGATAGACAGCAAGCGCGTGGGCCACCTACCACTGTCCGAGTGACACACA 1849
1850 GGTGTACATCTTAGAAGACGACACAGTCCAGTGTGACCTGTACAGCTGTACAGTCCCTGCG 1909
1850 AGTGTTTTATCTTCCCAATGACTCTATCCATTGTGAGAGAGAACTGTACCAATCGGCA 1909
1910 AGCCCTGGAAGACCAAGCTGCATCGACACGACGATGGAATGGAACCTTGCGAGACAAAA 1969
1910 GAGCGTGGAGGACCAATAGGCATACATTTGACAAAGAGATTGAAGCTCTGCAAGATAAAA 1969
1970 TTAAGAACCTGAGGGAAGTCCGAGGTCACTGTAAGAAAAAGCGGCAGAGAAATGTGACT 2029
1970 TTAAGAAATTAAGAGAGTGAAGAGACATCTGAGAGAGAGAGCTGAGGAAATGAGCT 2029
2030 GTCACAAATACAGTACCAACACCGACCAAGAGCGCGCTCAAGC-----ACAGAGGCT 2083
2030 GCAGTAAACAAAGCTATTACATAAAGAGAAAGGTGTAAAAAAGCAAGAGAAATTAAGA 2089
2084 CCAGTCTGCATCTTTTCAGAGAG--GGCTGCAAGAGAGAGAGAGAGTGTGGCTGTTCG 2140
2090 GCCATCTTACCATTTCAGAGAGCTGCTCAGGAGTAGATAGCAAACTGCAACTTTTCA 2149
2141 GGGAGCAGAGCGCAAGAAACTCCGCAAGTGTCTAAGCGCTTCGCAAGCAA---CG 2197
2150 AGGAGAACACCGTAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2209
2198 ACAGTGCAGCATGCCAGGCTCACTGCTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2257
2210 AAGAGTGCAGCTGCTGGGCTCACTTGTCTACGATGACAAACCACTGGCAGAGAGAG 2269
2258 CGGCTTTCTGAGACACTGGGCGCTTCTGCTGCTGACAGCGCCCAACATAACAGCTACT 2317
2270 CCGGTTCTGGAACCTGGGATCTTTCTGTGCTTGCAAGAGTTCTAACATAACACTACT 2329
2318 GGTGCATAGAGACCATCAATGAGACTCAATTTCTCTTCTGTGATTTGCAACTGGCT 2377
2330 GGTGTTTGGCTACAGTTAATGAGACGCAATTTCTTCTGTGAGTTGTCTAGTGGCT 2389
2378 TCCTAGAGTACTTTGATCTCAACACAGACCCCTACCAGCTGTGATGAATGCAATGAACACAC 2437
2390 TTTTGGAGTATTTTGTATGATGAATACAGATCTTTATCAGCTCAACATAAGTGCACAGG 2449
2438 TGGACAGGAGTCTCTCAACAGCTACAGCTCAAGTCTATGAGAGTGGAGAGCTGCAAGG 2497
2450 TAGAACGAGGCAATTTGAATCAGCTACAGCTACAGCTACAACTAATGAGGCTCAGAACTGTCAAG 2509
2498 GTTACAGAGAGTGTAAACCCCGGACTCGAAACATGAGCTGAGCTTAAAGATGGAGAA 2557
2510 GATATAAGCAGTGAACCCCAAGACCTAAGAAATCTTGAATGTGAAATTAAGATGGAGAA 2569
2558 GCTATGAGCAATACAG 2573
2570 GCTATGACCTACAG 2585

RESULT 5

US-09-773-426A-4
; Sequence 4, Application US/09773426A
; Patent No. 6534302
; GENERAL INFORMATION:
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Williamson, Mark
; APPLICANT: Tsia, Fong-ying

APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el
; FILE REFERENCE: 35800/208398(SB00-79
; CURRENT APPLICATION NUMBER: US/09/773,426A
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 09/495,823
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 4321
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (510)...(3125)
; NAME/KEY: misc_feature
; LOCATION: 4310
; OTHER INFORMATION: n = A,T,C or G
US-09-773-426A-4

Query Match 30.9%; Score 1040.4; DB 4; Length 4321;
Best Local Similarity 65.5%; Pred. No. 2.8e-257;
Matches 1660; Conservative 0; Mismatches 816; Indels 60; Gaps 7;
QY 98 TGTGGTGGTGAAGCTCGGCTTCTCTGTCGACCCACCGCTGAAAGCGCAGGTTTCAGAGGG 157
Db 559 TGTGGGAAGCCCTCTGTTCGACTGTGAGATCCCGAGGTTTCAGAGCGGATACAGCAGG 618
QY 158 ACOCAGGAAATCCGCCCCCAACATCATCTGTGTCTGACGAGCAGCAGGATGTGGAGC 217
Db 619 AAGCAAAAAACATCCGACCCCAACATTTCTTGTCTTACCGATGATCAAGATGTGGAGC 678
QY 218 TGGTTCATGTCAGGTGATGAACAAGACCGCGGCATCATGGAGCAGGCGGCGGCACCT 277
Db 679 TGGGTCTCTGCACTGATGAACAAACGAGAGAGATTATGAAACATGSGGGGCGCACCT 738
QY 278 TCATCAAGCGCTTCGTGACCAACCCATGTCGTGTCGCCCTCAACGCTCTCTCATCTCTCAGT 337
Db 739 TCATCAATGCTTTGTGACTACACCCATGTCGTGCGCGTCACGCTCTCTCATGCTCACCG 798
QY 338 GCAAGTACGTCCCAACCAACACACCTACACCAACATGAGAACTGCTCTCGCCCTCTCT 397
Db 799 GGAAGTATGTGCATCAATCAATGTCTACCAACAAACGAGAACTGCTCTTCCCTCTCGT 858
QY 398 GGCAGGCAACAGCAGAGCGCGCACTTTGCGCTGTACCTCAATAGACATGCGGTACCGGA 457
Db 859 GGCAGGCAATGATGAGCTCGGACTTTTGTCTGTATATCTTAACAACACTGGGCTACAGAA 918
QY 458 CAGCTTTCTTGGGAGTATCTTAATGAATACAAACGCTCTACGTGCCACCGGCTGGA 517
Db 919 CAGCTTTTGGGAAATACCTCAATGAATATATGAGCTACATCCCTCTGGGTGGC 978
QY 518 AGGAGTGGGTGGGACTCTTTAAAACTCCCGCTTTTATTAACACAGCTGTGTGGAAAG 577
Db 979 GAGAATGCTTGGATTAAATCAAGAAATTCCTGCTCTTATTAATTAACCTGTTGTGCAATG 1038
QY 578 GGTGAAAGAAAGCAGCGCTCCGACTACTCCAGAGATACCTCAGAGCTCATACCA 637
Db 1039 GCATCAAGAAAGCATGGAATTTGATTTATGCAAGGACTACTTTCAGAGCTTAATCACTA 1098
QY 638 ATGACAGCGTGAGCTTCTTCCGACGCTCCAAAGAGATGTACCCGCAACAGGCCAGTCTCA 697
Db 1099 ACAGAGCATTAATTAATTTCAAAATGTCTAAGAGATGATATCCCATAGGCCGTTATGA 1158
QY 698 TGGTCATCAGCATGAGCGCCCCCAACGCGCTGAGGATTCAGGCCCAATATTCAGGCC 757
Db 1159 TGGTCATCAGCATGAGCGCCCCCAACGCGCTGAGGACTCAGGCCCACTAGTTTCTAAAC 1218
QY 758 TCTTCCCAACGCGCTCAGCAGATCAGCGCGAGCTACAACTACGCGCCCAACCGGACA 817
Db 1219 TGTACCCCAATGCTTCCCAACACATAACTCTTAGTTAATTAATGACCAATATGGATA 1278

QY 818 AACCTGGATCATGCGCTACACGGGCGCCCATGAAGCCCATCCACATGGAATTCACCAACA 877
Db 1279 AACCTGGATCATGCGCTACACGGGCGCCCATGAAGCCCATCCACATGGAATTCACCAACA 1338
QY 878 TGCTCCAGCGGAAGCGCTTGCAGACCCCTCATCTCGGTGGACGACTCCATGGAGCGATTT 937
Db 1339 TTCTACAGCGCAAGGCTCCAGACTTTGATGTCTAGTGGATGATTTCTGTGGAGAGGCTGT 1398
QY 938 ACAACATGCTGGTTGAGACGGCGAGCTGGACACACACGTACATCGTATACACCGCGGACC 997
Db 1399 ATAACATGCTCGTGGAGACGGCGAGCTGGAGATATTTACATCATTTACACCGCGGACC 1458
QY 998 AGGTTTACACATCGGCCAGTTTGGCTGGTGAAGGGAATTCATGCGCATATGATGTTTG 1057
Db 1459 ATGGTTTACCATATGGCGAGTTTGGACTGGTCAAGGGGGAATTCATGCGCATATGATGTTTG 1518
QY 1058 ACATCAGGCTCCGTTTCTACGTGAGGGGCCCCAAAGTGGAGCGCGCTGTCTGAATCCCG 1117
Db 1519 ATATTCTGTGCTTTTATTTCTGTGGTCCAAAGTGTAGAACCGAGATCAATAGTCCAC 1578
QY 1118 ACATGCTCTCAACATGTGACCTGGCCCCCAACATCTCGACATTCAGGCTGGACATAC 1177
Db 1579 AGATGCTTCAACATGTGATTTGGCCCCCAACATCTCGATATTTGCTGGGCTCGACACAC 1638
QY 1178 CTGGGATATGACGGGAATCCATCTCAAGCTCTGGACACGGAGCGCGGCTGATC 1237
Db 1639 CTCCTGATGTGACGGCACTCTGCTCAAACTTCTGGAACCCAGAAAGCCAGGTAAAC 1698
QY 1238 GGTTCCTCACTGAAAAGAGATGAGGGCTCTGGCGGAGCTCTCTTGTGGAGAGAGCA 1297
Db 1699 GGTTCGGAACAAAGAGAGGCAAAATTTGGCGGTGATACATTTCTAGTGGAGAGGCA 1758
QY 1298 AGCTCTACACAGAGACACATGACAGGTGGAGCGCCAGAGAGAACTTTCTGCCCA 1357
Db 1759 AATTTCTAGTGAAGAGGAATCCACAGATATCCAAAGATTCACAGCTCAAAATCACTTTGCCA 1818
QY 1358 AGTACAGCGCTGTGAAGACCTGTGTGACGGGTGTGAGTACACAGCGCGGTGTGAGCAGC 1417
Db 1819 AATATGAAGGTCAAGACTATCCAGCAGCGGACAGTACACAGACCTGTGAAACAC 1878
QY 1418 TGGGACAGAGTGGCAGTGTGTGGAGGACGCAACGGGGAAGTGAAGTGTGATAGTGA 1477
Db 1879 CGGGGCGAAGTGGCAATGCAATGAGGATACATCTGGCAAGCTTCAAAATTCACAAAGTGA 1938
QY 1478 AGGGCCCCATCGGCTGGGCTGCTCAGCTCCGGCAGAGCACCGGAACTCTACGCTCGCGGT 1525
Db 1939 AAGGACCCAGTACCTGCTCAGCTCCGGCAGAGCACCGGAACTCTACGCTCGCGGT 1998
QY 1526 CCAAGTACTAGGGCAGG-----GCAGCGAGCGCTGCACCTGTGACAGCGGGGACT 1576
Db 1999 TCCATGACAAAGACAAAGAGTGCAGTTGTAGGGATCTGGTTACCGTGCACAGCAGAGCC 2058
QY 1577 ACAAGCTCAGCTGCGCGGACCGGGAAGAACTTTCAGAGAGATGACAGGCCAGCT 1636
Db 2059 AAAGAAAGAGTCAACGGCAATTTCTTGAGAAACCAAGGGGACTCCAAAGTACAAAGCCAGAT 2118
QY 1637 ATGTCCGAGTCCGCTCCATCCGCTCAGTGGCATCGAGGTGACGGCAGGGGTGTACACG 1696
Db 2119 TTGTCCATCTCGGACAGACAGTTCCTGTCTGTGCGTGAATTTGAAGGTGAATATATGACA 2178
QY 1697 TAGGCTGGGTGATGCCG-----CCAGCCCCGAAACCTCCACCAAGCGGC 1741
Db 2179 TAAATCTGGAAGAGAAAGAAATGCAAGTGTGCAACCAAGAAACATTCCTAAGGTC 2238
QY 1742 ACTGGCCAGGGGCCCCCTCAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACTG 1801
Db 2239 ATGATGAAGGCCAACAGGGGCCAAGAGATCTCCAGGCTTCAGTGGTGGCAACAGGGGCA 2298
QY 1802 GAGGCTTTCCGACTACTCAGCGCC-----AACCCCATTAAGGTGACATC 1849
Db 2299 GGATCTGGCAGATAGCAGCAACGCGTGGGCCCCACCTTACCCTGTCCGAGTGACACACA 2358

QY 1850 GGCTGTACATCTTAGAGAACGACACAGTGTCCAGTGTGACCTGGACCTGTCAAGTCCCTGC 1909
Db 2359 AGTGTTTATTTCTTCCCATGACTCTATCCATTTGTGAGAGAGACTGTACCATCGGCCA 2418
QY 1910 AGGCTTGAAGAACACACAGCTGCATCGACACACAGATTTGAACCCCTGCGAACAACAA 1969
Db 2419 GAGCGTGAAGAGGACCAATAGGCCATACATTTGACAAAGAGATTGAAGCTCTGCAAGATAAA 2478
QY 1970 TTAAGAACCTTAGGGAGTCCGAGGTCCACTGAAGAAAAAGCGCGCAGAGATGTGACT 2029
Db 2479 TTAAGAAATTAAGAGAGTGAAGGACATCTGAAGAGAGAGAGCTTGAGGAATGTAGCT 2538
QY 2030 GTCACAAATCAGCTACACACCCAGCACAAAGCGCGCTCAAGC-----ACAGAGGCT 2083
Db 2539 GCAGTAAACAAAGCTATTACAATAAGAGAGAAAGGTGTAAGAAAGCAAGAGAAATTAAGA 2598
QY 2084 CCAGTCTGCATCCCTTTCAGAGAG--GGCTCCAGAGAGAGACAGGTGTGGCTTGTGC 2140
Db 2599 GCATCTTCAACCATTCAAGAGGCTGCTCAGGAAGTAGATAGCAAACTGCAACTTTTCA 2658
QY 2141 GGGAGCAGAGCGCAAGAAAGAACTCCGCAAGTGTCTAAGCGCTCTGCAGAAACAA---CG 2197
Db 2659 AGGAGAACCAACCGTAGGAGAGAGAGAGAGAGAGAGCGGACAGAGAGAGGGGG 2718
QY 2198 ACAGTGCAGCATGCCAGGCTCAGTGTCTTCCACCAACAGCAACAGCACTGCGACAGG 2257
Db 2719 AAGAGTGCAGCGCTGCTGCTGCTTCCGCAATGCTGCTTCCGCAATGACAAACCACTGCGCAGACAG 2778
QY 2258 CGCTTTTCTGGACATCGGGGCTTTCTGTGCTGCACAGCGCCCAACATAACACGTACT 2317
Db 2779 CCGCTTCTGGAACCTGGGATCTTTCTGTGCTGCAGAGTTCTAACATTAACACCTACT 2838
QY 2318 GGTGCTAGAGGACCTCAATGAGACTCAAAATTTCTTCTGTGTAATTTGCAATTTGGCT 2377
Db 2839 GGTGTTGGCTCAGTAAATGAGAGCGCAATTTCTTCTGTGAGTTTGTACTGGCT 2898
QY 2378 TCTCAGTACTTGTATCTCAACAGACCCCTTACAGCTGATGATGACGTGACGTGACACAC 2437
Db 2899 TTTTGGAGTATTTGATATGAATACAGATCCCTTATCAGCTCAAAATACAGTGCACACG 2958
QY 2438 TGGACAGGATGCTCTCAACAGCTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGG 2497
Db 2959 TAGAACGAGCAATTTGAATCAGCTACAGCTACAGCTAATGAGCTCAGAAGCTGTCAAG 3018
QY 2498 GTTACAGCAGTGTAAACCCCGGACTCGAAACATGCACTGGACTTAAAGTGGAGGA 2557
Db 3019 GATATAAGCAGTGCAACCCCAAGACCTTAAGATCTTGTGTAATAAAGATGGAGGA 3078
QY 2558 GCTATGACCAATACAG 2573
Db 3079 GCTATGACCTACACAG 3094

RESULT 6

US-09-668-673B-15
; Sequence 15, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 4834
; TYPE: DNA

ORGANISM: Homo sapiens
US-09-668-673B-15

Query Match 29.8%; Score 1002.4; DB 4; Length 4834;
Best Local Similarity 65.7%; Pred. No. 1.7e-247;
Matches 1595; Conservative 0; Mismatches 771; Indels 60; Gaps 7;

QY 208 GATGTGGAGCTGGTTCCATGAGGTGATGAACAAGACCCGCGCATCATGAGAGCAGGC 267
DB 1 GATGTGGAGCTGGGTCCCTGCAAGTCATGAACAACGAGAAAGATTATGGAACATGGG 60
QY 268 GGGGCGCACTTCATCAACCCCTTCGTGACCAACCCATGTGTGCTCCCTCAAGCTCTCC 327
DB 61 GGGGCGCACCTTCATCAATGCCCTTGATGACACCCATGTGTGCTCCCTCAAGCTCTCC 120
QY 328 ATCCCTCACTGGCAAGTACCTCCACACACACACACCTACACCAACAAAGAGATGTCTCC 387
DB 121 ATGCTCACCGGAAGTATGTGCAATCAATCAATGTCTACACCAACAGAGAACTGTCTCT 180
QY 388 TGGCCCTCTCGGAGGACACGACGAGAGCCGACCTTTGGCGTGTACCTCAATAGCACT 447
DB 181 TCCCTCTCGTGGCAGGCCATGATGAGCTCGGACTTTTGTGTATATCTTAAACACT 240
QY 448 GGTACCGACAGCTTTCTTCGGGAAGTATCTTAATGAATACACGGCTCTCTAAGTGCCA 507
DB 241 GGTACAGAACAGCCCTTTTGGAAATACCTCAATGAATATAATGGCAGCTACATCCC 300
QY 508 CCGGCTGGAAGAGTGGGTGCGACTCTCTTAAAACTCCCGCTTTTATAACTACACGCTG 567
DB 301 CTGTGGTGGGAGATGGCTTGGATTATCAAGAACTCTCGTCTTATAATTAACCTGTT 360
QY 568 TGTGGAACCGGGTGAAGAAAGACGCTCCGACTTCCAGAGATACCTCAAGATTAACACAG 627
DB 361 TGTGCAATGGCATCAAGAAAGATGATGATTTGATGCAAGAGACTACTTCAAGAC 420
QY 628 CTCATCAACCAATGACAGCTGAGCTCTTCCGAGTCCCAAGATATACCGACAGG 687
DB 421 TTAATCACTAACAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
QY 688 CCAAGTCTCATGATCATGACCATGAGCCGCCACCGCCCTTGAAGATTCAGCCGCCAA 747
DB 481 CCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 748 TATTCAGGCTCTTCCCAACGCTCTGACATCATGCGCGAGTCAACTAGCGGCC 807
DB 541 TTTTCTAAACTGTACCCCAATGCTTCCCAACATTAATCTCTAGTTTAACTATGACCA 600
QY 808 AACCCGGAACAACTGTGATCATGCGCTACAGGGGCCCATGAAGCCCATCCACATGGA 867
DB 601 AATATGATATAACACTGGATTATGAGTACACAGGACCAATGCTGCCCATCCACATGGA 660
QY 868 TTCACCAATGCTCCAGCGGAGCGCTTGACAGCCCTCATGTCTGGTGAAGCTCCATG 927
DB 661 TTTTACAAACATTTACAGCGCAAAAGGCTTCCAGACTTTTGTATGTAGTGATGATGCT 720
QY 928 GAGACGATTTACAACTGTGTTGAGACGGCGAGCTGGACACACGCTACATCGTATAC 987
DB 721 GAGAGCTGTATAACATGCTCTGTGAGACGGGGAGCTGGAGATTAATCATATTAC 780
QY 988 ACCGCGACACAGGTTACATACATGCGCCAGTTTGGCTGTGTGAAGGAAATTCATGCCA 1047
DB 781 ACCGCGACCATGTTTACCATAATTTGGCAGTTTGGAGCTGTGTGAAGGAAATTCATGCCA 840
QY 1048 TATGAGTTTGAATCAGGTTCCGTTCTACGTGAGGGGCCCAAGTGAAGCGCGCTGT 1107
DB 841 TATGATTTTGAATATGCTGTGCTCTTTTATTTATTTATTTATTTATTTATTTATTTAT 900
QY 1108 CTGAATCCCACTCTCTCTCAACATGATGACCTTGGCCCCCACCACATCTCTGGAATTCAGGC 1167
DB 901 ATAGTCCCAAGATGCTTCTCAACTTGAATTTGGCTTTGGCCCCCACCACATCTCTGCTGG 960
QY 1168 CTGACATACCTCGGATATGAGCGGAAATCCATCTCTCAAGCTCTGGAACAGCGCGG 1227

DB 961 CTCGACACACCTCTCTGATGTGGAGCGCAAGTGTCTCTCAAACTTCTGGACCCAGAAAG 1020
QY 1228 CCGGTGAATCGGTTTCACTTTGAAAGAAAGATAGGGTCTGGCGGAGCTCTCTTTTGGTG 1287
DB 1021 CCAGGTAAACAGGTTTCCAAACAAACAGAAAGGCCAAATTTGGCGTGAATACATTTCTAGTG 1080
QY 1288 GAGAGAGGCAAGCTGTCTACCAAGAGAGACAATGACAAGGTGAGCCGCCAGAGAGAGAAC 1347
DB 1081 GAAAGAGGCNAATTTCTAGTAAGAGAGAAATCCAGCAAGATATCCACAGTCAAAAT 1140
QY 1348 TTTCTGCCCAAGTACCAAGGTGTGAAGACCTGTGTGACGCTGTGAGTACAGAGCGCG 1407
DB 1141 CACTTCCCAAAATATGACCGGTTCAAAGAACTATGCCAGAGCCAGGTACAGAGAGCC 1200
QY 1408 TGTGACAGCTGGGACAGAAAGTGGCAGTGTGGAGAGCGCCACGCGGAGGTGAAGCTG 1467
DB 1201 TGTGAACACCGGGGCGAAGTGGCAATGATGAGATATCATCTGGCAAGTTCGAAT 1260
QY 1468 CATAGTGAAGGGCCCCATGCGGTGGGC-----GGCAGCAGAGCCCTCTCC 1515
DB 1261 CACAAGTGAAGGACCCAGTGACCTGTCTACAGTCCGGCAGAGACGCGGAACCTCTAC 1320
QY 1516 AACTCGTCCCAAGTACTACGGGAGG-----GCAGCGAGCCCTGCACCTGTGAC 1566
DB 1321 GCTCGCGGTTCCATGACAAAGAGAGTCAACGGCAATTTCTTGAGAAACAGGGGACTCCAAAGTAC 1440
QY 1567 AGCGGGGACTTACAAGCTCAGCCTGCGCGAGCGCCGAAAAAACTCTTCAAGAAAGATAC 1626
DB 1381 AGCAGAAGCCAAAGAAAGAGTCAACGGCAATTTCTTGAGAAACAGGGGACTCCAAAGTAC 1440
QY 1627 AAGCCAGCTATGTCCGAGTGTCTCCATCCGCTCAGTGGCCATCGAGGTGAGCGCAGG 1686
DB 1441 AAGCCAGATTTTCCATCTCGGACAGACCTTCTTGTCCGTGCAATTTGAAGGTGAA 1500
QY 1687 GTGTACACAGTGGCTGGTGTATGCGCG-----CCAGCCCGCAAAACCTC 1731
DB 1501 ATATATGACATAAATCTGGAAGAGAGAAATTCAGAGTTTGCACCAAGAAACATT 1560
QY 1732 ACCAAGCGGACTGCGCCAGGGGCCCTTGAGGACCAAGATGACAAGATGGTGGGAGCTTC 1791
DB 1561 GCTAAGCGTCAATGATGAAGGCCACAGGGGCCAAGAGATCTCCAGGCTTCCAGTGTGCG 1620
QY 1792 AGTGGCACTGGAGGCTTCCCGACTACTCAGCGCG-----AACCCCAATAA 1839
DB 1621 AACAGGGCAGGATGCTGGCAGATAGCAGCAAGCGCTGGGCGCCACCTACGACTGCCA 1680
QY 1840 GTGACACATCGGTGTACATCTCTAGAGAACGACACAGTCCAGTGAACCTGAGACCTGTAC 1899
DB 1681 GTGACACACAAAGTGTATTTCTTCCCAATGACTCTATCCATTTGTGAGAGAGAACTGTAC 1740
QY 1900 AAGTCCCTCGAGCCCTGGAAGAACCAACAGCTGCATCGACCAAGAGATTGAAACCCCTG 1959
DB 1741 CAATCGGCGAGCGGTGGAGGAGCCATAAGGAGTACATTGAAGAGATTTGAAGCTCTG 1800
QY 1960 CAGAACAAATTAAGAACCTGAGGAGTCCGAGGTCACTTGAAGAAAAGCGCGCAGAA 2019
DB 1801 CAAGATAAAATTAAGAAATTTAAGAGAGTGAAGGACATCTGAAGAGAGAAAGCGCTGAG 1860
QY 2020 GAATGTGACTGTCAAAATCAGCTACACACCCAGCACAAGGCGCTCTCAAGCACAGA 2079
DB 1861 GAATGTGACTGTCAAGTAAACAAAGCTATTATCAATAAAGAGAAAGGTGTAAAGAAAGAGAG 1920
QY 2080 GGTCTCAG-----TCTGCATCTTTTCAAGAAAG-----GGCTGCAAGAGAGAGCAAGGTG 2130
DB 1921 AAATTAAGAGCCATCTTCAACCAATTCAGAGGAGTGTCTCAGGAAGTAGATAGCAAACTG 1980
QY 2131 TGGCTGTGGGAGCAGAGAGCGCAAGAAAGAACTCCGCAAGCTGTCTAAGCGCTGCGAG 2190
DB 1981 CAACTTTTCAAGGAGAACCAACCGTAGGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 2040
QY 2191 AACAC-----GACACGTGACAGCATGCCAGGCTTCAAGTCTTCAACCCACACACACAGCAC 2247
DB 2041 AGGAGGGGAAAGGTGACAGCTTGGCTTCACTTGTCTTCAAGCATGACAAACACAC 2100

QY 2248 TGSCACAGCGCGCTTTCTGGACACTGGGGCCTTTCTGTGCTGCACACGAGCCCAACAT 2307
Db 2101 TGSCACAGCGCGCTTTCTGGACACTGGGGCCTTTCTGTGCTGCACAGGTTCTAACAT 2160
QY 2308 AACAGTACTGGTGCATGAGGACCAATCAATGAGACTCAAAATTCCTCTTCTGTGAATTT 2367
Db 2161 AACACCTACTGGTGTGTTGGGTACAGTTAATGAGAGCGCATAAATTTCTTTCTGTGAGTTT 2220
QY 2368 GCAACTGGCTTCTAGAGTACTTTGATCTCAACACAGACCCCTACAGCTGATGAATGCA 2427
Db 2221 GCTACTGGCTTTTGGAGTATTTGATATGATATAGATCTCTTATCAGCTCACAAATACA 2280
QY 2428 GTGAACACACTGACAGGAGATTCCTCAACAGCTACAGCTACAGCTCATGAGCTGAGG 2487
Db 2281 GTGCACACGCTAGACAGGAGCAATTTGATCAGCTACAGCTACAGCTAATGAGCTCAGA 2340
QY 2488 AGCTCAAGGGTTACAAGCAGGTAAACCCCGGACTCGAATCATGAGCTGGACTTAA 2547
Db 2341 AGCTGTCAAGGATATAGCAGGTGCAACCCCAAGACCTAAGAAATCTTGTATGTTGGAATAA 2400
QY 2548 GATGAGGAAGCTATGAGCAATACAG 2573
Db 2401 GATGAGGAAGCTATGACCTACACAG 2426

RESULT 7

US-09-668-673B-13

; Sequence 13, Application US/09668673B

; Patent No. 6562956

; GENERAL INFORMATION:

; APPLICANT: Emerson, Charles P

; APPLICANT: Dhoot, Gurtej K

; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF

; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR

; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS

; FILE REFERENCE: PENN-0733

; CURRENT APPLICATION NUMBER: US/09/668,673B

; CURRENT FILING DATE: 2000-09-22

; PRIOR FILING DATE: 1999-09-23

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 1367

; TYPE: DNA

; ORGANISM: Mus sp.

US-09-668-673B-13

Query Match 29.8%; Score 1001.6; DB 4; Length 1367;
Best Local Similarity 86.2%; Pred. No. 1.5e-247;
Matches 1181; Conservative 3; Mismatches 162; Indels 24; Gaps 6;
QY 1335 CCAGGAGGAGAACTTTCTGCCAAGTACACGGTGTGAAGACCTGTGTACGGTGTCTGA 1394
Db 1 CCAGGAGGAGAACTTTCTGCCAAGTACACGGTGTGAAGACCTGTGTACGGAGCTGA 60
QY 1395 GTACCAGACGGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTGAGGACCGCCAGG 1454
Db 61 GTACCAGACAGCATGCAAGCAGCTGGGACAGAGTGGCAGTGTGTGAGGACCGCTTCTGG 120
QY 1455 GAAGCTGAAGCTGCATAAGTGAAGGGCCCATCGGCT-----GGGCGGCAGCAG 1505
Db 121 CACGCTGAAGCTGCATAAATGTAAAGGCCCATCGGCTTTGGTGGCGGCTGGCAGCAG 180
QY 1506 AGCCCTCTCCAACTCTGTCGACAGTACTACGGGAGGCGGAGCGCTGACCTGTGA 1565
Db 181 AGCCCTCTCCAACTCTGTCGACAGTATGACGGCCAGAGCAGCGAGGCTGCACTGTGA 240
QY 1566 CA-----GCGGGGACTACAACTCAGCTCAGCTGCGCGGACCGGAAACAACTCTTCAA 1616
Db 241 CAGTGGCGGTGGAGGGGACTACAACTGGGCTGGCTGGAGCGCCG-----TAAGCTCTTTAA 297

QY 1617 GAAGAGTACAAGCGCGAGCTATGTCCGAGTGCCTCCATCCGCTCAGTGGCCATCAGGT 1676
Db 298 GAAAAGTATAAGACAGCTATGCCGGAAACCGCTCCATCCGTTCCGTGGCCATCAGGT 357
QY 1677 GAGCGGAGGCTGTACACGCTAGGCTGGGTGATGCGCGCCAGCGCCCGAAACCTCACCAA 1736
Db 358 GAGCGGTGAGATATACCGTAGGCTTGGATATCTGTGCTCAGCCCCGACCTTAGCAA 417
QY 1737 GCGGCACTGCGCAGGCGCCCTGAGACCAAGATGACAAAGATGTTGGGAGCTTCACTGG 1796
Db 418 GCGGCACTGCGCAGGCGCCCTGAGACCAAGATGACAAAGATGTTGGGAGCTTCACTGG 477
QY 1797 CACTGAGGCGCTTCCCGAATCTACTCAGCGCCCAACCCATTAAGATGACACATCGGTGCTA 1856
Db 478 TACTGTGGCTTCCAGATTAATCTGCCCCCAATCCCATCAAGTGAACCATCGGTGCTA 537
QY 1857 CATCTTAGAAGACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGAGGCTG 1916
Db 538 CATCTTAGAAGATGACACAGTCCAGTGTGACCTGTACAAGTCCCTGAGGCTTGG 597
QY 1917 GAAAGACCAACAGCTGCATCGACACGAGATGAAACCCCTGCAGAACAAAATTAAGAA 1976
Db 598 GAAAGACCAACAGCTGCATCGACCTGAGATGAAACCCCTGCAGAACAAAATTAAGAA 557
QY 1977 CTTGAGGGAAGTCCGAGGTCACTGAAAGAAAGCGGCCAGAGAAATGTGACTGTACAA 2036
Db 658 CTTGAGGAAGTCAAGGGTCACTGAAAGAAAGCGGACCGGAAGAAATGTGACTGTCAATA 717
QY 2037 AATCAGCTACCAACCCAGCACAAGGCGCCCTCAAGCAGACAGGCTCCAGTCTGCATCC 2096
Db 718 AATCAGTTACCAACCAACAAAGGCGCTCTCAAGCACAAGGCTCCAGCTGCGACCC 777
QY 2097 TTTTCAAGAAAGGCTGCAAGAGAGGACAAAGTGTGGTGTTCGCGGAGCAGAAGCGCAA 2156
Db 778 TTTTCAAGAAAGGCTGCAAGAGAGGACAAAGTGTGGTGTTCGCGGAGCAGAAGCGCAA 836
QY 2157 GAAGAACTCCGAGAGCTCTCAAGGCTCGAAGAACAAAGCAGCAGTGCAGATGCCAGG 2216
Db 837 GAAGAACTCCGAGAGCTCTCAAGGCTCGAAGAACAAAGCAGCAGTGCAGATGCCAGG 894
QY 2217 CTTCAAGTCTTCAACCCAGCACAAGCAGCTGCGGCGGCTTCTTGGACACTGGG 2276
Db 895 CTTCAAGTCTTCAACCCAGCACAAGCAGCTGCGGCGGCTTCTTGGACACTGGG 954
QY 2277 GCCTTTCTGTGCTGCACAGCGCCCAACAAATAACGTAACGTAACGTAACGTAACGTAAC 2336
Db 955 GCCTTTCTGTGCTGCACAGCGCCCAACAAATAACGTAACGTAACGTAACGTAACGTAAC 1014
QY 2337 TGAGACTCACAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCT 2396
Db 1015 TGAGACTCACAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCT 1074
QY 2397 CAACACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGAACAGGATGTCTCAA 2456
Db 1075 CAGTACAGACCCCTACAGCTGATGAACGCGGTGAACACACTGGAACAGGAGCTCTTAA 1134
QY 2457 CCAGCTACAGTACAGTCTGAGAGCTGAGAGCTGCAAGGGTTACAAGCAGTGTAAACC 2516
Db 1135 CCAACTGCACTGAGCTCATGAGAGCTAAGGAGCTGTAAAGGCTCAACAGCAGTGAACCC 1194
QY 2517 CCGGACTCCGAAACATGGACCTTGAAGTGAAGTGAAGGAGCTATGAGCAATACAGCA 2576
Db 1195 CCGGACTCCGAAACATGGACCTTGAAGTGAAGTGAAGGAGCTATGAGCAATACAGCA 1254
QY 2577 GTTTCAGGCTGAAAGTGGCCAGAAATGAAGAGACTTTCTTCCAAATCACTGGGCAACT 2636
Db 1255 GTTTCAGGCTGAAAGTGGCCAGAAATGAAGAGACTTTCTTCCAAATCACTGGGCAACT 1314
QY 2637 GTGGAGAGCTGGGAGGTTAAGAAACAAACAGAGTGCACCTCCAAAC 2686
Db 1315 ATGGGAAGTGGGAGGCTAAGCGCCATAGAGAGAGAACTCCAAAC 1364

RESULT 8

US-09-810-347-1
; Sequence 1, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL001169
; CURRENT APPLICATION NUMBER: US/09/810,347
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Human
US-09-810-347-1

Query Match 17.2%; Score 578.2; DB 4; Length 1799;
Best Local Similarity 73.1%; Pred. No. 1.1e-138;
Matches 742; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
QY 98 TGGTGGTGGAGCTGGCTTCTCTGTCGACACCAACCGCTGAAAGCAGGTTTCAGAGG 157
Db 156 TGGTGGTGGAGCTGGCTTCTCTGTCGACACCAACCGCTGAAAGCAGGTTTCAGAGG 215
QY 158 ACCGAGGAACATCCGCCCAACATCATCTCTGTCGACGACGACGACGAGGATGGAGC 217
Db 216 AACGAAAAAATCCGACCCCAACATATTCTTGTGCTTACCGATGATCAAGATGGAGC 275
QY 218 TGGGTTCATGCGAGTGATGAACAGACCGCGGCGATCATGAGCGAGCGCGGCGCACT 277
Db 276 TGGGTTCATGCGAGTGATGAACAGACCGCGGCGATCATGAGCGAGCGCGGCGCACT 335
QY 278 TCATCAACGCTTCTGTCGACACCAACCGCTGTCGCTCTCCCTCAGCTCTCTCCCTCAGT 337
Db 336 TCATCAACGCTTCTGTCGACACCAACCGCTGTCGCTCTCCCTCAGCTCTCTCCCTCAGT 395
QY 338 GCAAGTAGCTCCCAACCAACCAACCTTACCAACCAACCAACCAACCAACCAACCAAC 397
Db 396 GGAAGTAGTGTGCAACATCAACATGTCTACACCAACCAACCAACCAACCAACCAAC 455
QY 398 GGCAGGACACGACGAGCGGCGACCTTTCGCTGTACCTCAATAGCACTGCTACCGA 457
Db 456 GGCAGGACACGAGCTGCGGACTTTCGCTGTATCTTACCAACCACTGCTACCGA 515
QY 458 CAGCTTCTTCGGAAGATATCTTAAATGAATACCAACCGCTCTACGTCGCGCTCTCT 517
Db 516 CAGCTTCTTCGGAAGATATCTTAAATGAATATATGCGAGCTACATCCCTCTGGTGGC 575
QY 518 AGAGTGGTGGGACCTTAAACCTCCGCTTTTAACTACAGCTGTGTGCGAGC 577
Db 576 GAGATGGTGGGATTAATCAAGATTTCTGCTTCTTAAATTAATTAATTAATTAATTA 635
QY 578 GGGTGAAGAAACACGCGCTCCGACTTCTCCAGGATTAACCTCAGACCTCATCAACCA 637
Db 636 GCATCAAGAAAGCATGGATTGATTAATGCAAGGACTTCTCAGACTTAATCACTA 695
QY 638 ATGACAGCTGAGCTTCTCCGACGCTCCAGAGATGTACCGGACAGGCGAGCTCTCA 697
Db 696 ACAGAGCATTAATTAATCTTCAAAATGTCTAAGAGATGTATCCCATAGGCGCTTATGA 755
QY 698 TGGTCTACGCCATGACGCGCCCTCGAGGATTCAGGATTCAGCCCAACATATTCAGGCG 757
Db 756 TGGTGTACGCCATGAGCTGCGCCCTCGAGGATTCAGGATTCAGCCCAACAGTTTCTA 815
QY 758 TCTTCCCAACGAGCTTACGACATCAACGCGGAGCTAAGCTACGCGCCCAACCGGACA 817
Db 816 TGTACCCCAATGCTTCCCAACATCAATCTAGTTAATTAATTAATTAATTAATTAATTA 875
QY 818 AACACTGGATCATGGCTACACGCGGCCCCATGAAGCCCCATCCACATGGAAATTCACCA 877

Db 876 AACACTGGATTCAGCTACACAGGACCAATGCTGCCATCCACATGAATTTACCAACA 935
QY 878 TGGTCCAGCGGAGCGCTTGGAGACCTCATGTCGGTGGAGCGACTCCATGGAGACGATTT 937
Db 936 TTTCTACAGCGCAAAAGGCTCCAGACTTTGATGTGTCAGTGGATGATTTCTGTGGAGGCTGT 995
QY 938 ACAACATGCTGGTTGACACGCGGAGCTGGACAAACAGTACATCGTATACACCGCGGACC 997
Db 996 ATAACATGCTGCTGGAGACGCGGGAGCTGGAGATCTTACATCATTTACACCGCGGACC 1055
QY 998 AGGTTTACCATCGGCCAGTTTGGCTGTGTAAGAGGAATCCATGCCATATGAGTTTG 1057
Db 1056 ATGTTTACCATATTTGGCAGTTTGGACTGCTCAAGGGGAATCCATGCCATATGACTTTG 1115
QY 1058 ACATCAGGTCCTGCTTCTACGTGAGGGGCCCCACGTGGAGCGGCTGCTCTGAA 1112
Db 1116 ATATCGTGGCTTTTATTTTATTCGTGCTCAAGTGTAGAACCCAGGATCAATGTA 1170
RESULT 9
US-09-404-879A-315/c
; Sequence 315, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-315
Query Match 13.1%; Score 441; DB 4; Length 441;
Best Local Similarity 100.0%; Pred. No. 9.6e-104;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2895 CAAAAATGCAATTTTTCGTATCAAAAAGTCAACCTTCCCGCAGAGCTCACAA 2954
Db 441 CAAAAATGCAATTTTTCGTATCAAAAAGTCAACCTTCCCGCAGAGCTCACAA 382
QY 2955 AGGAAACGAGAGAGAGAGAGAGATTTTCTTGGAAATTTCTCCCAAGGCGAAAG 3014
Db 381 AGGAAACGAGAGAGAGAGAGAGATTTTCTTGGAAATTTCTCCCAAGGCGAAAG 332
QY 3015 TCATGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTCTTAAATCTCTTATTTCT 3074
Db 321 TCATGGAATTTTAAATCATAGGGGAAAAGCAGTCTCTTCTTAAATCTCTTATTTCT 262
QY 3075 TGGTTCTCACAAGAGAGAGAGAGAGAGAGATTTTCTTGGAAATTTCTCCCAAGGCGAAAG 3134
Db 261 TGGTTCTCACAAGAGAGAGAGAGAGAGATTTTCTTGGAAATTTCTCCCAAGGCGAAAG 202
QY 3135 CAGTGCAGAGACGTTTGGACAAATGAGTCAAGTACAGCAAAAAGAGATGACATTTACCTAGCAC 3194
Db 201 CAGTGCAGAGACGTTTGGACAAATGAGTCAAGTACAGCAAAAAGAGATGACATTTACCTAGCAC 142
QY 3195 TATAAACCTTGGTTCCTCTGAAGAACTGCTTCAATGATATATGACTATTTTACAT 3254
Db 141 TATAAACCTTGGTTCCTCTGAAGAACTGCTTCAATGATATATGACTATTTTACAT 82
QY 3255 GTATCAACATGGGAACTTTTAGGGGAACCTTAATAAGAAATCCCAATTTTCAGGAGTGT 3314
Db 81 GTATCAACATGGGAACTTTTAGGGGAACCTTAATAAGAAATCCCAATTTTCAGGAGTGT 22

QY 3315 GGTGTCAATAAACGCTCTGTG 3335
Db 21 GGTGTCAATAAACGCTCTGTG 1

RESULT 10

US-09-668-673B-10
; Sequence 10, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-668-673B-10

Query Match 11.3%; Score 381.6; DB 4; Length 494;
Best Local Similarity 89.5%; Pred. No. 1.9e-88;
Matches 444; Conservative 0; Mismatches 49; Indels 3; Gaps 3;
QY 2003 AGAAAGCGGCCAGAGAGATGTGACTGTCAAAATCAGCTACACACCAGCAAAAG 2062
Db 1 AGAAGAGCGACCGGAAGATGTGCTGCCATAAAATCAGTTACACAGCCCAACACAAAG 60
QY 2063 GCGGCTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGG 2122
Db 61 GCGGCTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGG 120
QY 2123 ACAAGTGTGGTGTGGGAGCAGAGAGCGCAAGAACTCGCAAGTGTCTCAAGC 2182
Db 121 ACAAGTGTGGTGTGGGAGCAGAGAGCGCAAGAACTCGCAAGTGTCTCAAGC 178
QY 2183 GCTGCAGAACACACACAGCTGCACATGCCAGGCTCACTGCTTACCCACAGCAACC 2242
Db 179 GCTGCAGAACACACATACGTGACATGCC-GGCTCACGTGCTTACCCACAGCAACC 237
QY 2243 AGCACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGCTGCTGCAACGCGCA 2302
Db 238 ACCACTGGCAGACGGCGCACTCTGACGCTGGGCGCTTTCTGCTGCTGCAACGCGCA 297
QY 2303 ACAATAACACGTACTGTGTCATGAGGACATCATGAGACTCAAAATTCCTTCTGTG 2362
Db 298 ACAATAACACGTACTGTGTCATGAGGACATCATGAGACTCAAAATTCCTTCTGTG 357
QY 2363 AATTGGCACTGGCTTCTAGAGTACTTTGATCTCAACACAGACCCCTACCACTGATGA 2422
Db 358 AATTGGCACTGGCTTCTAGAGTACTTTGATCTCAACACAGACCCCTACCACTGATGA 417
QY 2423 ATGCAGTGAACACACTGACAGGAGTCTTCAACAGCTACAGTACAGTCTATGGAGC 2482
Db 418 AGCGGTTGAACACACTGACAGGAGTCTTCAACAGCTACAGTACAGTCTATGGAGC 477
QY 2483 TGAGGAGTGAAGGG 2498
Db 478 TAAGGAGTGAAGG 493

RESULT 11

US-09-668-673B-9
; Sequence 9, Application US/09668673B
; Patent No. 6562956

; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-668-673B-9

Query Match 11.2%; Score 375.8; DB 4; Length 466;
Best Local Similarity 90.6%; Pred. No. 5.6e-87;
Matches 423; Conservative 0; Mismatches 42; Indels 2; Gaps 2;
QY 1885 GACCTGACCTGTACAGTCCCTGACGGCTGGAAGACCAACAGCTGCATCGACCC 1944
Db 2 GACCTGACCTGTACAGTCCCTGACGGCTTGGAAAGACCAACAGCTGCATCGACCCAT 61
QY 1945 GAGATTGAAACCTCGAGAACCAAAATTAAGAACCTTGAGGAAGTCCGAGGTCACTGAAG 2004
Db 62 GAGATCGAAACCTCGAGAACCAAAATTAAGAACCTTGAGGAAGTCCGAGGTCACTGAAG 121
QY 2005 AAAAAGCGGCCAGAGATGTGACTGTCAAAATCAGCTACACACCAGCAAAAGC 2064
Db 122 AAGAGCGACCGGAAGATGTGACTGCCATAAAATCAGTTACCAAGCAACACAAAGC 181
QY 2065 CGCTCAAGCACAGAGGCTCCAGTCTGCACTCTTCAGGAAGGGCTGCAAGAGAGGAC 2124
Db 182 CGTCTCAGCACAAAGCTTCCAGCTTCCAGCTTCCAGTCTGAGGAGGCTGAGGAGAGGAC 241
QY 2125 AAGGTGTGCTGTGGGAGCAGAGCGCAAGAACTCCGCAAGTGTCTCAAGCGC 2184
Db 242 AAGGTGTGCTGTGGGAGCAGAGCGCAAGAACTCCGCAAGTGTCTCAAGCGC 299
QY 2185 CTGCAGAACACACACAGCTGCACATGCCAGGCTCAGGCTTACCCACAGCAACCG 2244
Db 300 CTGCAGAACACACATACGTGACATGCCAGGCTCAGGCTTACCCACAGCAACCG 359
QY 2245 CACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGCTGCTGCAACGCGCAAC 2304
Db 360 CACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGCTGCTGCAACGCGCAAC 419
QY 2305 AATAACACGTACTGTGTCATGAGGACCAATCATGAGACTCAAAATTT 2351
Db 420 AATAACACGTACTGTGTCATGAGGACCAATCATGAGACTCAAAATTT 466

RESULT 12

US-09-484-970B-70
; Sequence 70, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 70
; LENGTH: 3871
; TYPE: DNA

Db 181 ACAGCTGATGAACCGGGTGAACACACTGGACAGGACGCTCTTAACCAACTGCACGTGC 240
QY 2471 AGCTCATGAGCTGAGGAGCTGCAAGGGTTTACAAGCAGTGTAAACCCCGGACTTCGAACA 2530
Db 241 AGCTCATGAGCTAAGGAGCTGTAAAGGCTACAAGCAGTGTAAACCCCGGACCCCAACA 300
QY 2531 TGGACCTGGGACTTAAAGATGGAAGGAGCTATGAGCAATACAGGAGTTCAGCGTCGAA 2590
Db 301 TGGACCTGGGCTTAGAGACGAGGAGGAGCTATGAGCAATACAGGAGTTCAGCGTCGAA 360
QY 2591 AGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGGNAGCTGGG 2650
Db 361 AATGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAGCTATGGGAAGTTCGG 420
QY 2651 AAGGTTAAGAAACAACAGAGGTTGGACCTCCAAAAAC 2686
Db 421 AAGGCTAAGCGGCCATAGAGAGAGAACTCCAAAAAC 456

RESULT 14
US-09-668-673B-8/c
; Sequence 8, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-668-673B-8

Query Match 10.3%; Score 347.4; DB 4; Length 538;
Best Local Similarity 82.1%; Pred. No. 1.2e-79;
Matches 444; Conservative 0; Mismatches 76; Indels 21; Gaps 3;

QY 1335 CCAGGAGAGAACTTTCGCCCAAGTACAGCGTGTGAAGCACTGTGTACAGCGTGTGA 1394
Db 538 CCAGGAGAGAACTTTCGCCCAAGTACAGCGGCTGAAGGACCTGTGTACAGCGTGTGA 479
QY 1395 GTACCAGACGGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTGAGGACGCGTTCGG 1454
Db 478 GTACCAGACAGCATGCGAACAGCTGGGCGACAGTGGCAGTGTGTGAGGACGCGTTCGG 419
QY 1455 GAAGCTGAAGCTGCANATAGTCAAGGGCCCATCGGCT-----GGCGCGCAGCAG 1505
Db 418 GACGCTGAAGCTGCACAAATGTAAAGGCCCATCGGCTTTGGTGGCGCGCTGGCAGCAG 359
QY 1506 AGCCCTCTCCAACTCGTGGCCCAAGTACTAGCGGACGAGGACGAGGCTTGCACCTGTGA 1565
Db 358 AGCCCTCTCCAACTCGTGGCCCAAGTATGACGGCCAGACGAGGAGGCTTGCAGCTGTGA 299
QY 1566 CA-----GCGGGGACTACAAGCTCAGCGCTGGCGGACGCGGAAAAAACTTTCAA 1616
Db 298 CAGTGGCGGTGGAGGGGACTACAAGCTGGGCTTGGCTGGAGCGCG---TAAGCTCTTAA 242
QY 1617 GAAAGAGTACAGGCGCAGCTATTCGCGAGTTCGCTTCATCGCTCAGTGGCCATTCAGGTT 1676
Db 241 GAAAAAGTATAAGACCAAGCTATGCCCGGAACCGCTCCATCCGTTCCGTGGCCATCGAGGT 182
QY 1677 GGACGGCAGGCTGTACCACTAGGCTGGGCTGGGTGTAGCGCCCGACCGCCAACTCCAA 1736

Db 181 GGACGCTGAGATATACCACGTAGGCTGGATACTGTGCTCAGCCCCGCAACCTTAGCAA 122
QY 1737 GCGGCACTGCGCAGGGGCCCCCTGAGACCAAGATGACAGGATGGTGGGACTTTCAGTGG 1796
Db 121 GCGGCACTGCTCAGGGGCCCCCTGAGACCAAGATGACAGGATGGTGGGCAAGTTTTCAGTGG 62
QY 1797 CACTGGAGGCTTCCCGACTACTCAGCGGCAACCCCATTTAAAGTGCACATCGGTGCTA 1856
Db 61 TACTGGTGGCTTCCAGATTATTCTGCCCCCAATCCCATCAAGTGCACCATCGGTGCTA 2
QY 1857 C 1857
Db 1 C 1

RESULT 15
US-09-668-673B-11
; Sequence 11, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-668-673B-11

Query Match 8.8%; Score 297.4; DB 4; Length 436;
Best Local Similarity 83.4%; Pred. No. 7.6e-67;
Matches 366; Conservative 0; Mismatches 61; Indels 12; Gaps 2;

QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGGGGCGAGGCGGAGGCTGCACC 1560
Db 1 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGGGGCGAGGCGGAGGCTGCACC 60
QY 1561 TGTGACA-----GCGGGGACTACAAGCTCAGCGCTGGCGGACGCGGAAAAAACTC 1611
Db 61 TGTGACAGTGGCGGTGGAGGGGACTACAAGCTGGGCTGGTGGACGCG---TAAGCTC 117
QY 1612 TTCAGAAGAACTACAGCGCAGCTATGTCCGAGTCCGCTCATCGCTCAGTGGCCATC 1671
Db 118 TTTAAGAAAAAGTATAAGACCAAGCTATGCCGGAACCGCTCCATCCGTTCCGTGGCCATC 177
QY 1672 GAGGTGGACGCGGAGGCTGTACCACTAGCGCTGGGTGTATGCCGCCAGCCCGGAAACCTC 1731
Db 178 GAGGTGGACGCTGAGATATACCACTAGCTTGGATCTGTGCCTCAGCCCGCCGAACTT 237
QY 1732 ACCAAGCGGCACTGCGCAGGGGCCCTTGAAGCAAGATGACAAAGATGGTGGGCACTTC 1791
Db 238 AGCAAGCGGCACTGCGCAGGGGCCCTTGAAGCAAGATGACAAAGATGGTGGGCACTTC 297
QY 1792 AGTGGCACTGGAGGCTTCCCGACTACTCAGCGGCAACCCCATTTAAAGTGCACATCGG 1851
Db 298 AGTGGTACTGGTGGCTTCCAGATTATTCTGCCCCCAATCCCATCAAGTGCACCATCGG 357
QY 1852 TGCTACATCTTAGAAGACGACACAGTCCAGTGTGACCTTGGACCTGTACAAAGTCCCTGCAG 1911
Db 358 TGCTACATCTTAGAAGATGACACAGTCCAGTGGCAGCTTGGACCTGTACAAAGTCCCTGCAG 417
QY 1912 GCTCGAAGACCAACAGC 1930
Db 418 GCTTGAAGACCAACAGC 436

Search completed: February 14, 2004, 19:42:07
Job time : 235 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 13:53:07 ; Search time 6194 Seconds
(without alignments)
13195.989 Million cell updates/sec

Title: US-10-025-966A-4

Perfect score: 3363

Sequence: 1 gactccgcgcatcccaaga.....aaagaaaaaa 3363

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Maximum DB seq length: 2000000000
Listing first 45 summaries

Database :

EST:

1: em_estba.*

2: em_esthum.*

3: em_estmu.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_estc.*

9: gb_estl.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gssl.*

29: gb_gss2.*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2093.6	62.3	3704	11 AK036685	AK036685 Mus muscu
2	2093.6	62.3	3759	11 AK034712	AK034712 Mus muscu
3	2092	62.2	3687	11 AK028874	AK028874 Mus muscu
4	2092	62.2	3734	11 AK081643	AK081643 Mus muscu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	1430.8	42.5	2685	11 AK049170	AK049170 Mus muscu
6	1062.4	31.6	4564	11 AK045002	AK045002 Mus muscu
7	1062.4	31.6	4513	11 AK078756	AK078756 Mus muscu
8	1060.8	31.5	4428	11 AK028285	AK028285 Mus muscu
9	856.4	25.5	1803	11 AK008108	AK008108 Mus muscu
10	838.2	24.9	1201	9 AL537332	AL537332 Mus muscu
C	829.4	24.7	1201	9 AL574439	AL574439 Mus muscu
11	829.4	24.7	1201	9 AL574439	AL574439 Mus muscu
12	812.8	24.2	948	13 BU956086	BU956086 AGENCOURT
13	799	23.7	1066	13 BU956086	BU956086 AGENCOURT
14	798.6	23.7	885	13 BU956086	BU956086 AGENCOURT
15	791.4	23.5	904	13 BU956086	BU956086 AGENCOURT
16	788.4	23.4	861	14 CD106296	CD106296 AGENCOURT
17	786.4	23.4	933	13 BX370010	BX370010 AGENCOURT
18	777.6	23.1	916	12 BU911672	BU911672 AGENCOURT
19	760.8	22.6	899	14 CA489747	CA489747 AGENCOURT
20	754.8	22.4	843	13 BU858200	BU858200 AGENCOURT
21	749.2	22.3	874	13 BU858200	BU858200 AGENCOURT
22	748.6	22.3	2703	11 AK037389	AK037389 Mus muscu
23	742.8	22.1	899	13 BU858200	BU858200 AGENCOURT
24	725	21.6	856	13 BU858200	BU858200 AGENCOURT
25	722.8	21.5	967	10 BG386102	BG386102 AGENCOURT
C	721.4	21.5	890	13 BX411283	BX411283 AGENCOURT
C	719	21.4	730	14 CB850913	CB850913 UI-CF-EN1
28	713.2	21.2	941	12 BU910212	BU910212 AGENCOURT
29	709.2	21.1	826	12 BU908810	BU908810 AGENCOURT
30	704.4	20.9	916	13 BU171062	BU171062 AGENCOURT
31	700.6	20.8	982	13 BU930803	BU930803 AGENCOURT
C	697.8	20.7	736	13 BU448044	BU448044 UI-H-EU1
C	697	20.7	728	12 BM985096	BM985096 UI-CF-RC1
34	696.2	20.7	739	13 BU448116	BU448116 UI-H-EU1
35	695.8	20.7	981	13 BU645715	BU645715 AGENCOURT
36	694.2	20.6	880	13 BU927919	BU927919 AGENCOURT
C	693.6	20.6	1201	9 AL566767	AL566767 Mus muscu
C	693	20.6	705	13 BU446043	BU446043 UI-H-EU1
C	692.8	20.6	1201	13 BU461722	BU461722 AGENCOURT
C	692.6	20.6	1045	10 BG036300	BG036300 AGENCOURT
C	689.4	20.5	703	13 BU181056	BU181056 UI-H-EU0
C	688.8	20.5	703	12 BM981713	BM981713 UI-CF-EN1
42	685	20.4	753	10 BF530592	BF530592 AGENCOURT
43	685	20.4	755	10 BG248293	BG248293 AGENCOURT
44	682.8	20.3	755	10 BG248293	BG248293 AGENCOURT
45	676.8	20.1	704	12 BM930827	BM930827 UI-E-EU1

ALIGNMENTS

RESULT 1

AK036685

LOCUS

DEFINITION

AK036685 3704 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male bone cDNA, RIKEN full-length enriched
library, clone:9830162M14 product:weakly similar to
N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert
sequence.

ACCESSION

AK036685

VERSION

AK036685.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

Db 585 ACCTTGGCGGTGTATCTCAACAGCACAGGCTACCGGACAGCTTCTTCTCGGAAAAATACCTC 644
QY 481 AATGAATAACACGGCTCTACTGTCACACCGCGCTGGGAAGAGTGGGTGCGACTCTCTTAAA 540
Db 645 AATGAGTACACAGGCTCATAGTGGCGCGCGCTGGGAAGAGTGGGTGCGGCTTACTTAA 704
QY 541 AATCCCGCTTTTATACTACACGCTGTGTGGAAACGGGTGAAAGAAAAAGCAGCGCTCC 600
Db 705 AATCCCGCTTTTATACTACACACTCTGCGCGGAATGGGGTGAAGGAGAAACATGGCTCA 764
QY 601 GACTACTCTCAAGGATTAACCTCACAGACCTCATCAACCAATGACACGCTGAGCTTCTCCG 660
Db 765 GACTACTCCAGGATTAACCTCACGATCTCATCAACCAATGACAGCTGAGCTTCTCCGA 824
QY 661 AGTCCAGAAAGATGTACCCGCACAGGCGAGTCTCATGTGTCATCAGCCATGAGGCCGCC 720
Db 825 ACATCCAGAAAGATGTACCCACACAGGCGCGTGTCTCATGGTTCATCAGCCACGCGGCTCC 884
QY 721 CACGCGCTGAGGATTCAGCCCAATATTCAGGCTCTTCCCAACGATCTCAGCAC 780
Db 885 CATGCCCCGAGGACTCAGCACCCGAGTACTCAGCGGTCTTCCCAATGGCTCCGAGCAC 944
QY 781 ATCAGCGGAGCTCAACTACGCGCCCAACCCGACAAACACTGGATCATGCGCTACAG 840
Db 945 ATCAGCGGAGTACAACTATGCAACCCCAACCCAGACAGCATTTGGATCATGCGCTACAG 1004
QY 841 GGGCCCATGAAGCCCATCCACATGGAATTCACCAATGTCTCCAGCGGAAGCGTTGCGAG 900
Db 1005 GGACCCATGAAGCCCATTCACATGGAATTCACCAATGTCTACCAACGAAACGCTACAG 1064
QY 901 ACCCTCATGTGCGGAGGACTCCATGAGACGATTTACAACTGCTGTTGAGAGCGGC 960
Db 1065 ACCCTCATGTCTGGATGATCTCATGAGAGCATCTATGACATGCTGGTGGAGACGGG 1124
QY 961 GAGTGGCAACACAGTACATCGTATACACCGCGACCAACGCTTACCAACATCGGCAGTTT 1020
Db 1125 GAGTGGCAACACAGTACATCGTATACACCGCGACCAACGCTTACCAACATGCGCAGTTT 1184
QY 1021 GGCCTGGTGAAGGGAATCATCCCATATGAGTTTGCATCAGGTCGCGTCTACGTG 1080
Db 1185 GGCCTGGTGAAGGGAAGTCTATGCGGTATGAATTCGACATCAGATGCCGTTCTACGTG 1244
QY 1081 AGGGCCCCAACGTTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCAACATGACCTG 1140
Db 1245 AGGGCCCCAACGTTGGAAGCTGCTCTCTGAAACCCACATTTGCTCTCAACATGACCTG 1304
QY 1141 GCGCCACCATCTGGACATTTGAGGCGCTGACATACCTGCGGATATGGAACGGGAATCC 1200
Db 1305 GCGCCACCATCTGGATATGCTGGACTGGACATCCCTGACAGACATGGAACGGGAAGTCT 1364
QY 1201 ATCTCTAAGCTGTGGACACGGAGCGCGGTGAATCGGTTTCACTTGAAGAAAGAGATG 1260
Db 1365 ATCTCAAACTACTGGACTGAGAGCGGCGAGTGAACCGGTTCCACTTGAAGAAAGAGCTG 1424
QY 1261 AGGCTCTGGCGGACTCTCTTCTGGTGGAGAGCGCAAGCTGTCTACAAAGAGAGACAAT 1320
Db 1425 AGGCTCTGGCGGACTCTCTTCTGGTGGAGAGCGCAAACTGTCTCCACAAAGAGCGGCT 1484
QY 1321 GACAGGTGAGCGCGGAGGAGACTTTCTGCCCAAGTACACGCTGTGAGAGGAGTGGCAGTGTG 1380
Db 1485 GACAAAGTGAATGCCGAGGAGAACTTCTGCCCAAGTACACGCTGTGAGAGGAGGACCTG 1544
QY 1381 TGTGAGCGTGTGAGTACAGACGCGGCTGTGAGCAGTGGGACAGAAAGTGGCAGTGTG 1440
Db 1545 TGTGAGCGAGTGTGAGTACAGACAGCATGCGAACAGCTGGGGCAGAAAGTGGCAGTGTG 1604
QY 1441 GAGGACCGCAAGGAGAGCTGAAGTGAAGTGAAGGCGCCCATGCGGCT----- 1493
Db 1605 GAGGACGCTTCTGGGAGGCTGAAGCTGCACAAATGTAAAGGCCCATGCGGCTTGTGGC 1664
QY 1494 --GGGCGGCGACGAGCGCTCTCAACCTCGTGGCCCAAGTACTACGGGCGAGGCGAG 1551
Db 1665 GCGGCTGGCACGAGCGCTCTCAACCTGGTGGCCCAAGTATGAGCGCCAGAGCGGAG 1724

QY 1552 GCCTGACCTGTGACA-----GCGGGGACTACAAAGCTCAGCTCGCGGCGGACGCGG 1602
Db 1725 GCCTGACCTGTGACAGTGGCGGTGAGGGGACTACAAAGCTGGGCTGTGCTGGACGCG- 1783
QY 1603 AAAAACTCTTCAAGAAGAAGTACAAAGGCCAGCTATGTCCGCACTCGCTTCCATCCGCTCA 1662
Db 1784 --TAAGCTCTTTAAGAAAAAGTATTAAGACCACTATATGCGCGGAACCGCTCCATCGTTCC 1841
QY 1663 GTGGCCATCGAGGTGGACGGGAGGTGTACCACTGAGGCTGGGTGTATGCCGCCAGGCC 1722
Db 1842 GTGGCCATCGAGGTGGACGGGTGAGATATACCACTGAGGCTTGGATCTGTGCTCAGGCC 1901
QY 1723 CGAAACCTCACCAAGCGGCACTGGCCAGGGGCCCTCTGAGGACCAAGATGAACAAGATGGT 1782
Db 1902 CGCAACCTTAGCAAGCGCACTGGCCAGGGGCCCTGAAGACCAAGATGAACAAGATGGT 1961
QY 1783 GGGGACTTCACTGAGGCACTGGAGGCCCTTCCGACTACTCAGCGGCCCAACCCATTAAGTG 1842
Db 1962 GGCAGTTTCACTGAGTCTGAGTGGCTTCCAGATTAATCTGCCCCCAATCCATCAAGTG 2021
QY 1843 ACACATCGGTGTACATCTTAGAGAAACACACAGTCCAGTGTGACTGCACTGCACTGTACAG 1902
Db 2022 ACCCATCGGTGTACATCTTGAAGATGACACAGTCCAGTGCAGTTCGACTTGGACCTGTACAG 2081
QY 1903 TCCCTGCGGCTGGAAGAAGACCAAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 1962
Db 2082 TCCCTGCGGCTTGGAAAGAGACCAAGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2141
QY 1963 AACAAATTAAGAACTCTGAGGAGTCCGAGGTGCACTGAGAAAGAAAGCGGCCAGAGAA 2022
Db 2142 AACAAATTAAGAACTCTGAGAGTCTGAGGAGTCACTGAAAGAAAGCGGACCGGAGAA 2201
QY 2023 TGTGACTGTCAAAATCAGCTTACCAACCCAGCACAAAGGCGCGCTCAAGCACAGAGGC 2082
Db 2202 TGTGACTGTCAATAGATCAGTTACCAAGCCAAACAAAGGCGCTCTCAAGCAAAAGGC 2261
QY 2083 TCCAGTGTGCATCTTTTTCAGGAGGCGCTGAGAGAGAGGCAAGGTGTGCTGTGCGG 2142
Db 2262 TCCAGCTGTCACTCTTTTTCAGGAGGCTCTGAGGAGAGGCAAGGTGTGCTGTGCGG 2321
QY 2143 GAGCAGAGCGCAGAAAGAACTCCGCAAGCTGTCAAGCGCTTGCAGAAACAAACGACAGC 2202
Db 2322 GAGCAGAAACGCAAGAAAGAACTCCGCAAGCTGTCAAAACGGCTGCAGAAACAAACGATAGC 2381
QY 2203 TGCAGATGCGCAGGCGCTCACTGCTTACCCACAGCAACAGAGCACTGGCAGCGGCGCT 2262
Db 2382 TGCAGCATGCGCGGCTTCACTGCTTACCCACAGCAACCACTGCGCAGCGCGGCA 2441
QY 2263 TTTCTGACACTGGGGCTTTCTGTGCTGCACTGCAAGCGCCACCAATTAACAGTCTGTGTC 2322
Db 2442 CTCTGAGCGTGGGGCGGTTCTGTGCTGCACTGCAAGCGCCACCAATTAACAGTCTGTGTC 2501
QY 2323 ATGAGGACCATCAATGAGACTCACAAATTTCTCTTGTGAAATTTGCAACTTGGCTTCTCTA 2382
Db 2502 TTGAGGACCAATAATGAGACCCACAACTTCTCTTCTGCGAAATTTGCAACCGGCTTCTATA 2561
QY 2383 GAGTACTTTGATCTCAACACAGACCCCTTACAGCTGATGATGAGTGAACACACACTGAGC 2442
Db 2562 GAATACTTTGACCTCAGTACAGACCCCTTACAGCTGATGAAACGCGGTGAACACACTGAGC 2621
QY 2443 AGGGATGCTCTCAACAGCTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTAC 2502
Db 2622 AGGGAGCTCTTAAACCACTGCACTGCGAGCTCATGAGCTTAAGAGCTGTAAAGGCTAC 2681
QY 2503 AAGCAGTGTAAACCGGAGCTGAAACATGAGACCTGGACTTTAAAGATGAGGAGGCTAT 2562
Db 2682 AAGCAGTGTAAACCGGAGCTGAAACATGAGACCTGGGCTTTAGAGACGAGGAGCTAT 2741
QY 2563 GAGCAATACAGCAGTCTTTCAGCGTGAAGGTGGCCAGAAATGAAGAGACCTTCTTCCAAA 2622
Db 2742 GACCAATACAGCAGTCTTTCAGCGTGAAGGTGGCCAGAAATGAAGAGACCTTCTTCCAAA 2801

QY 2623 TCACTGGGCAACTGTGGGAAGCTGGGAAGTTAAGAAACACAGAGGTGGCACTCCAA 2682
 Db 2802 TCACTGGGACAGCTATGGGAAGCTGGGAAGCTAAGCGGCATAGAGAGGAACCTCC 2861
 QY 2683 AAACATAGAGGATCATCACTGACGACAGGCAATGAATAA 2722
 Db 2862 AAACACAGGGGCTGTGTGGTGTGCCAGGCCATGCAAAA 2901

RESULT 2
 AK034712
 LOCUS
 DEFINITION Mus musculus 12 days embryo embryonic body between diaphragm region
 and neck cDNA, RIKEN full-length enriched library, clone:9430027K19
 product:weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix
 coturnix], full insert sequence.

ACCESSION AK034712
 VERSION AK034712.1 GI:26330130
 KEYWORDS HTC; CAP trapper;
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schiraldi, L. M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmington, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J. H., Kotsuki, S.
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660

11217851

PUBMED
REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3759)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirokawa, T.,
 Hori, P., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-resgsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 A cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/.

FEATURES

source

1. .3759
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:9430027K19"
 /db_xref="taxon:10090"
 /clone="9430027K19"
 /tissue_type="embryonic body between diaphragm region and
 neck"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 160. .2886
 /notes="unnamed protein product; putative
 weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE
 [Coturnix coturnix] (SPTR|Q90XB6, evidence: FASTY,
 64.2%ID, 99.5%length, match=2601)"
 /codon_start=1
 /protein_id="BAC28804.1"
 /db_xref="GI:26330131"
 /translation="NANPGRGEGIPSSRHHGLCIQBEAKDQQTAPPGPLPLWLL
 INEAGAHFINAVTTCPCPSRSITLTKYVHNHTYNNENSCSSPQAOHESRTF
 STAGLSLAGSAFLSHKGRFORNIRNPILVLTDDQVDELGSMQVWKNKTR
 AYLNSTGYRTAPFGKYLNEVNGSYPPQKEMVGLKNRFYNTLCENGKXKHS
 DYSTDLTLINTSVSPFRFTSKMYPHPLVLSHAAPHGSPASQYSLFPNAS
 QHTPSYIAPNDPDRHINRYTGMKPIHMEFTNMLQRKQLTLMSSVDSMTIYDML
 VETGELDNTIILYTAGDHIQIGLVKSGMPYFDFRVPYFVGNVAGSLNPHI
 VLNIADPILDIADIDPADMGSKILKLSERFVNRFLKRLVRWDSFLVBERG
 KILKREGKQVNAENFLPKYQRYDKQAEYQACBQLGKQCVCDASGTLKXK
 KCKPMRGFGGGSRALSNLVPKQSSACSCDGGGDKYKGLAGRRKLPKKYK
 TSVARNRSIRSVIAEVDGRIYHVGDITVPQENLSKPHWPAPEDDDKGGFSGTG
 GLPDSAPNPIKVTHRCYILENDTVQCDLILYKSLQAWDKHLIDHEIETLQNKIK
 LREVRHLLKKKPECCDCHRSYTHSGHKLXKSSLPFRKRGKDEKQWLLREKQ
 RKCKLKLKLLQNNDTCSMPGLCTFTHDNHWQTAPLTWLGPFCACYSANNNTWCL
 RTINETHFLFCFEFTLYFDLSDPYQLMANVNTLDRVNLQVLMELRSCKG
 YKQCNPRTRNMDLGLRDGGSVEYQFQRRKWPENKMPSSKSLGLQWEGW"

CDS

BASE COUNT

962 a 1086 c 957 g 754 t

ORIGIN

[illegible]

QY	1021	GGCCTGGTGAAGCGGAAATCCATGCCATATGAGTTTGAACATCAGGGTCCCGTTCTACGTTG	1081
DB	1234	GGGCTGGTGAAGGGCAAGTCTATTCGCCGTATGAATTCGACATCAGAGTCCCGTTCTACGTTG	1293
QY	1081	AGGGGGCCCCAACCGTTGGAAGCGCGCTCTCTGAATCCCCACATCGTCTCTCAACATTGACCTTG	1140
DB	1294	AGGGGGCCCCAACCGTTGGAAGCTGGCTCTCTGAAACCCCCACATTTGTCTCAACATTGACCTTG	1353
QY	1141	GCGCCACCATCTTGACATTTGCAGCGCTTGGAATACCTTCCGGATATGACGCGGAATATCC	1200
DB	1354	GCGCCACCATCTTGATATTCGTGTGACTTGGACATCCCTGCGACATGGAACGGGAAGTCT	1413
QY	1201	ATCCTCAAGCTGTGGACACGGAGCGCGCGGTGAATTCGGTTTCACTTGAATAAAGAAGATG	1260
DB	1414	ATTCTCAAACTACTGGACTCAGAGCGGCCAGTGAACCGGTTCCACTTGAATAAAGAAGCTG	1473
QY	1261	AGGGTCTGGCGGACTCTCTTCTTGTTGGAGAGAGGCGAAGCTGCTACACAGAGAGACAAT	1320
DB	1474	AGGGTCTGGCGGACTCTCTTCTTGTTGGAGAGAGGCAAACTGCTCCACAAGAGGGAGGGT	1533
QY	1321	GACAAGTGTGACCCCGAGGAGGAGAACTTTCTTGCCCAAGTACACAGCGTGTGAAGGACCTTG	1380
DB	1534	GACAAAGTGAATGCCCGAGGAGGAACTTTCTTGCCCAAGTACACAGCGTGTGAAGGACCTTG	1593
QY	1381	TGTCAGGGTCTGAGTACACGACGGCGTGTGACAGCTGGGACAGNAGTGGCAGTGTGTG	1440
DB	1594	TGTCAGCGAGCTGAGTACACGACAGCACTGCGAAACAGCTGGGGCAGAAGTGTGGCAGTGTGTG	1653
QY	1441	GAGGACGCCACGGGGAAAGCTGAAGCTTGCAATAAGTGCAAGGGGCCCATGCGGGCT-----	1493
DB	1654	GAGGACGGCTCTTGGGACGCTGAAGCTGCAAAATGTAAAGGCCCATGCGGTTTGGTGGC	1713
QY	1494	--GGGGCGGACGACGAGCCCTCTCCAACTCTGTGCCCCAAGTACTACGGGCGAGGCGACGAG	1551
DB	1714	GGGGTGGGACGACAGCCCTCTCCAACTTGGTGCCCCAAGTACTACGGGCGAGGCGACGAG	1773
QY	1552	GCTTGACCTGTGACA-----GCGGGGACTACAAGCTCAGCGCTGGCCGCGACGCCGG	1602
DB	1774	GCCTGCAGCTGTGACAGTGGCGGTGGAGGGGACTACAACTGGGCGCTGGCTGGAACGCCG-	1832
QY	1603	AAAAAATCTTCAAGAGAAAGTACAAAGCCAGCTATGTCGCGAGTCTGCTCCATCCGCTCA	1662
DB	1833	--TAGCTCTTTAAGAAAGATATAGACCACTATGCCCGGAAACGCTTCCATCCGTTCC	1890
QY	1663	GTGGCCATCAGGTGGAACGGCAGGGTGTACCAAGTAGGCTGGGTGATGTCGCCCCAGGCC	1722
DB	1891	GTGGCCATCAGGTGGAACGGTGGATATACCAAGTATAGGCTTGGATACTGTGCCTCAGGCC	1950
QY	1723	CGAAACCTCACCAAGGGCACTTGGCCAGGGGCCCTGAGGACCAAGATGACAAAGATGGT	1782
DB	1951	CGCACTTTAGACGGCCACTTGGCCAGGGGCCCTGAAGACCAAGATGACAAAGATGGT	2010
QY	1783	GGGGACTTCAGTGGCACTGGAGGCCCTTCCCGACTACTCAGCGGCCAAACCCCATTAAGTGT	1842
DB	2011	GGCAGTTTTCAGTGGTACTTGGTGCCCTTCCAGATTATTTCTGGCCCCCAATCCCATCAAGTG	2070
QY	1843	ACACATCGGTGTACATCTGAGAAACGACACAGTCCAGTGTGACTCGACTGTATACAG	1902
DB	2071	ACCATCGGTGTACATCTTGGAAATGACACAGTTCAGTGGCATTTGAACTCTGTACAAAG	2130
QY	1903	TCCCTGCGGCTGGAAGACCAACAAGCTGCAATCGACCAAGAGATTGAACCTCTGAG	1962
DB	2131	TCCCTGCGGCTTGAAGAACCAACAAGCTGCAATCGACCAAGATGAGATGAAACCTCTGAG	2190
QY	1963	AACAAATTAAGAACTGAGGGAAGTCCGAGGTCACTTGAAGAAAAACGGCCAGAGAA	2022
DB	2191	AACAAATTAAGAACTTCGAGAGTTCAGGGTCACTTGAAGAAAGGACCGGAGAA	2250
QY	2023	TGTGACTGTCAAAATCAGTTTACCAACCCAGCAACAAAGGCCGCTCAAGCACAGAGGC	2082
DB	2251	TGTGACTGTCAATAGAAATCAGTTTACCAAGCAACCAAGAGGCCGCTCTCAAGCACAAAGGC	2310

QY	2083	TCAGTCTGCATCCTTTTCAGAGAGGCGCTCGAAGAGAGCAAGGCTGTGGCTGTGGCG	2142
Db	2311	TCAGGCTGACACCTTTTCAGGAGGCTGCGAGAGAGGCAAGGCTGTGGCTGTGGCG	2370
QY	2143	GACGAGAGCGCAAGAGAGAACTCCGCAAGCTGCTCAAGCGCTCGAGAACACGACAG	2202
Db	2371	GAGCAGAGAGCGCAAGAGAACTGCGCAAGCTGCTCAAGCGCTCGAGAACACGACAG	2430
QY	2203	TGCGATGCGCAGGCTCGCTGCTTCAACCCAGCAACACGACCTGGCGAGCGCGCT	2262
Db	2431	TGCGATGCGCAGGCTCGCTGCTTCAACCCAGCAACACGACCTGGCGAGCGCGCA	2490
QY	2263	TTCTGAGCACTGGGCGCTTCTGTGCTCGACGCGCCCAACAACTACTGTGCTGC	2322
Db	2491	CTCTGAGCGTGGGCGCTTCTGTGCTCGACGCGCCCAACAACTACTGTGCTGC	2550
QY	2323	ATGAGGACCATCAATGAGACTCACAAATTTCTCTCTGTGAAATTTGCAACTGGCTTCCCTA	2382
Db	2551	TTGAGGACCATCAATGAGACTCACAAATTTCTCTCTGTGAAATTTGCAACTGGCTTCCCTA	2610
QY	2383	GAGTACTTTGATCTCAACAGACCCCTACGAGCTGATGAATGCACTGACACTGGAC	2442
Db	2611	GAATACTTTGATCTCAACAGACCCCTACGAGCTGATGAATGCACTGACACTGGAC	2670
QY	2443	AGGAGTGTCTCAACAGAGCTACACGCTACGAGCTCATGAGCTGAGGAGCTGCAAGGTTAC	2502
Db	2671	AGGAGTGTCTCAACAGAGCTACACGCTACGAGCTCATGAGCTGAGGAGCTGCAAGGTTAC	2730
QY	2503	AGCAGTGTAAACCCCGGACTGCAACATGAGACCTGGGACTTAAAGATGGAGGAAGCTAT	2562
Db	2731	AAAGCAGTGAACCCCGGACTGCAACATGAGACCTGGGACTTAAAGATGGAGGAAGCTAT	2790
QY	2563	GAGCAATACAGGAGTTTTCAGGCTGCAAGAGTGGCAGAAATGAGAGAGCTTCTTCCAAA	2622
Db	2791	GAACATACAGGAGTTTTCAGGCTGCAAGAGTGGCAGAAATGAGAGAGCTTCTTCCAAA	2850
QY	2623	TCACTGGGCAACTGTGGGAGGCTGGGAGGTTAAGAAACACAGAGTGAGCTCCAA	2682
Db	2851	TCACTGGGCAACTGTGGGAGGTTGGGAGGTTAAGGCGCCATAGAGAGAGGAGCTCC	2910
QY	2683	AAACATAGAGGATCACTGACTGACAGCGCAATGAATAA	2722
Db	2911	AAACATAGAGGATCACTGACTGACAGCGCAATGAATAA	2950
RESULT 3			
LOCUS	AK028874		
DEFINITION	Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732465J15 product:weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert sequence.		
ACCESSION	AK028874		
VERSION	1	GI:26324821	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM			
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999).		
PUBMED	99279253		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Hashiwa, K., Fujiwaka, S., Inoue, K., Iogawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, D. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. P., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauer, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3687)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source	1..3687	/organism="Mus musculus"	QY	481	AATGAATACAAACGGCTCTCTACGTGCCACCCCGCTCGAAGAGAGTGGGTGGAGTCTCCTTAAA	540
		/mol_type="mRNA"	Db	624	AATGAGTACAAACGGCTCATACGTGCCCGCCCGCTGGAAGAGTGGGTGGCTTAAAG	583
		/strain="C57BL/6J"	QY	541	AACTCCCGCTTTTATAAATACACGTGTGTGGGAACGGGGTGAAGAAAGAACGCGCTCC	600
		/db_xref="PANTOM,DB:4732465J15"	Db	684	AACTCCCGCTTTTATAAATACACACTCTGCGCGGAATGGGGTGAAGAGAAACATGGCTCA	743
		/db_xref="taxon:10090"	QY	601	GACTACTCAAGGATTACCTCACAGACCTCATCAACATGACAGCGGTGAGCTTCTTCCGC	660
		/clone="4732465J15"	Db	744	GACTACTCAAGGATTACCTCACGGATCTCATCCGATCTCATCACCANTGACAGTGTGCTTCCGA	803
		/tissue_type="skin"	QY	661	AGCTCAAGAGATGATACCGGACAGGCCAGTCTCATGGTTCATCAGCCATGACAGCCGCC	720
		/clone_libs="RIKEN full-length enriched mouse cDNA library"	Db	804	ACATCAAGAAGATGATACCAACACAGGCCCGTGTCTCATGGTTCATCAGCCACGCGGCTCCC	863
		/dev_stage="10 days neonate"	QY	721	CACGGCCCTGAGGATTACAGCCGCCCAATATATACAGCTCTTCCCAAAACGATCTCAGCAC	780
		189..2816	Db	864	CATGCCCCGAGGACTCAGCACCCAGTACTACCGCTCTTCCCCAATGCGTCCGAGCAC	923
		/note="unnamed protein product; putative	QY	781	ATCACGCCGAGTACAACTACCGGCCCAACCCGGACAAACACTGGATCATCGCTTACACG	840
		weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE	Db	924	ATCACACCGAGTTACAACTATGTCACCCCAACCCAGCAAGCAATTGGATCATCGCTTACACG	983
		[Corumix coturnix] (SPR[Q90XB6, evidence: FASTY,	QY	841	GGGCCCCATGAGCCCATCCACATGGAATTCACCACTGCTCCAGCGGAAGGCTTGCAG	900
		64.2%id, 99.5%length, match=2601)"	Db	984	GGACCAATGAGCCCATTCACATGGAATTCACCACTGCTTACCAACGATGCTACCAACGCGCTACAT	1043
		/codon_start=1	QY	901	ACCTTCATGTCGGTGGAGACTCCATGAGAGACGATTTTCAACATGCTGCTGTTGAGACGGGC	960
		/protein_id="BAC26165.1"	Db	1044	ACCTTCATGTCGGTGGAGACTCCATGAGAGACGATTTTCAACATGCTGCTGTTGAGACGGGC	1103
		/db_xref="GI:26324822"	QY	961	GAGCTGGACACACGTATACATCGTATACACCCGACACACGCTTACCATCGGCGAGTTT	1020
		translating="G1:26324822"	Db	1104	GAGCTGGACACACGTATACATCGTATACACCCGACACACGCTTACCATCGGCGAGTTT	1163
			QY	1021	GGCTGGTGAAGGGAATCCATGCTATGAGTTTGACATCAGGCTCCCGTCTTCTACGTG	1080
			Db	1164	GGCTGGTGAAGGGAATCCATGCTATGAGTTTGACATCAGGCTCCCGTCTTCTACGTG	1223
			QY	1081	AGGGCCCCAACGCTGGGAAGCGGCTGTGTAATCCCACTCGCTCTCAACATTTGACCTG	1140
			Db	1224	AGGGCCCCAACGCTGGGAAGCGGCTGTGTAATCCCACTCGCTCTCAACATTTGACCTG	1283
			QY	1141	GCCCCCACCCTCTGGACATTCAGGCTGACATCTCGGATATGAGCGGGAATCC	1200
			Db	1284	GCCCCCACCCTCTGGACATTCAGGCTGACATCTCGGATATGAGCGGGAATCC	1343
			QY	1201	ATCTCTCAAGCTGCTGGACACGAGCGCGCGTGAATCGGTTTCTCACTTGAAGAAAGATG	1260
			Db	1344	ATCTCTCAAGCTGCTGGACACGAGCGCGCGTGAATCGGTTTCTCACTTGAAGAAAGATG	1403
			QY	1261	AGGGCTGGCGGAGTCTCTTCTGGTGGAGAGAGCGGCTGCTACACAGAGAGACAT	1320
			Db	1404	AGGGCTGGCGGAGTCTCTTCTGGTGGAGAGAGCGGCTGCTACACAGAGAGACAT	1463
			QY	1321	GACAGGTGGACGCCCGGAGGAGAACTTTCTGCCCAAGTACCAAGCGGTGTGAAGACCTG	1380
			Db	1464	GACAGGTGGACGCCCGGAGGAGAACTTTCTGCCCAAGTACCAAGCGGTGTGAAGACCTG	1523
			QY	1381	TGTCAAGCTGTGAGTACAGACCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG	1440
			Db	1524	TGTCAAGCTGTGAGTACAGACCGGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG	1583
			QY	1441	GAGGACGCCCGGAGGAGCTGAAGCTGCAAGTGAAGGCGGCCCTCCATCGGCT	1493
			Db	1584	GAGGACGCCCGGAGGAGCTGAAGCTGCAAGTGAAGGCGGCCCTCCATCGGCT	1643
			QY	1494	--GGCGGACGACAGAGCCCTCTCCAACTCTGTGCCCAAGTACTACGGGCGAGGCGAG	1551
			Db	1644	GGCGGACGACAGAGCCCTCTCCAACTCTGTGCCCAAGTACTACGGGCGAGGCGAG	1703
			QY	1552	GCCTGCACCTGTGACA-----GGGGGAGCTACAGCTCAGCTGCGCGGACGCGG	1602
BASE COUNT	946 a 1062 c 948 g 731 t					
ORIGIN						
Query Match	62.2%; Score 2092; DB 11; Length 3687;					
Best Local Similarity	86.1%; Pred. No. 0;					
Matches 2359; Conservative	0; Mismatches 360; Indels 21; Gaps 3;					
QY	1	GACTCCCGCATCCCAAGAGACACAGATCAGCAAAAGAAAGATGGGCCCGCCGAGC	60			
Db	144	GGCTCTGTCATCCAGGAGAGCAAGAGACAGCAGCAAGCCACGCGCAATGGCACCCCTGGC	203			
QY	61	CTCGTGTGCTGTGCTGTGCTCGCACTGTGTCTTCCCTGTGCTGTGAGCTCGGCTTC	120			
Db	204	CTGCCATATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	263			
QY	121	CTGTGCAACACCGCTGAAAGGAGGATTTTCAAGGAGACCGAGGAACATCCGCCCCCAAC	180			
Db	264	CTGTGCAACACCGCTGAAAGGAGGATTTTCAAGGAGACCGAGGAACATCCGCCCCCAAC	323			
QY	181	ATCATCTGGTGTGAGGAGACGAGATGTGGAGTGGGTTCCATGAGGTGATGAAC	240			
Db	324	ATCATCTGGTGTGAGGAGACGAGATGTGGAGTGGGTTCCATGAGGTGATGAAC	383			
QY	241	AAGACCGGCGCATCATGAGACGAGGCGGGCGCACTTCATCAACGCTTCTGTGACCA	300			
Db	384	AAGACCGGCGCATCATGAGACGAGGCGGGCGCACTTCATCAATGCTTCTGTGACCA	443			
QY	301	CCCATGTGCTCCCTCAGCTTCTCATCTCTCATCTGGCAAGTACGTCCCAACCAAC	360			
Db	444	CCCATGTGCTTCTCGCTGTGCTTCTCATCTGGCAAGTACGTCCCAACCAAC	503			
QY	361	ACCTACCAACCAATGAGACTGCTCTCGCTTCTGCGGAGACGAGCAGGAGCGCG	420			
Db	504	ACCTACCAACCAATGAGACTGCTCTCGCTTCTGCGGAGACGAGCAGGAGCGCG	563			
QY	421	ACCTTTGCGGTGATCAATAGACATGGCTACCGGACAGCTTTCTTTCGGGAAAGTATCT	480			
Db	564	ACCTTTGCGGTGATCAATAGACATGGCTACCGGACAGCTTTCTTTCGGGAAAGTATCT	523			

Db 1704 GCTCGAGCTGTGACAGTGGCGGTGGAGGGGAGTACAACTGGGCTGTGGAGCGCG- 1762
QY 1603 AAAAACTCTTCAAGAGAGTACAGGCGAGTATGTCGCGAGTGTGCTTCATCGGTCA 1662
Db 1763 --TAAGCTCTTTAAGAAAAAGTATAGGCGCAGTATGTCGCGAAGACCGTTCATCGTTCC 1820
QY 1663 GTGGCCATCGAGGTGGAGCGGCTGTACACAGTGTAGGCTGGGTGTAGCGCCAGCC 1722
Db 1821 GTGGCCATCGAGGTGGAGCGGCTGTACACAGTGTAGGCTGGGTGTAGCGCCAGCC 1880
QY 1723 CGAAACCTTCAACAGCGGCTGTGGCGAGGCGGCTGTAGGACCAAGATGCAAGAGTGGT 1782
Db 1881 CGCAACCTTGTAGCAAGCGGCTGTGGCGAGGCGGCTGTAGGACCAAGATGCAAGAGTGGT 1940
QY 1783 GGGGACTTCAGTGGGCTGTGGGCGGCTTCCGACTACTCAGCGCGCAACCCATTAAGTG 1842
Db 1941 GGCAGTTTCAGTGGGCTGTGGGCGGCTTCCGACTACTCAGCGCGCAACCCATTAAGTG 2000
QY 1843 ACACATCGGTGTACATCTCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAG 1902
Db 2001 ACCATCGGTGTACATCTCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAG 2060
QY 1903 TCCTTCAGGCTGTGAAGACACAGTGTGACCTGTGACCTGTGACCTGTGACCTGTGAC 1962
Db 2061 TCCTTCAGGCTGTGAAGACACAGTGTGACCTGTGACCTGTGACCTGTGACCTGTGAC 2120
QY 1963 AACAAATTAAGAACTGTAGGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAG 2022
Db 2121 AACAAATTAAGAACTGTAGGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAG 2180
QY 2023 TGTGCTGTCAAAATCAGTGTACACACGACGACCAAGGCGGCTGTGAGGAGGAGGAGGAG 2082
Db 2181 TGTGCTGTCAAAATCAGTGTACACACGACGACCAAGGCGGCTGTGAGGAGGAGGAGGAG 2240
QY 2083 TCCAGTGTGCTCTTTCAGGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2142
Db 2241 TCCAGTGTGCTCTTTCAGGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2300
QY 2143 GAGCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2202
Db 2301 GAGCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2360
QY 2203 TCCAGTGTGCTCTTTCAGGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2262
Db 2361 TCCAGTGTGCTCTTTCAGGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2420
QY 2263 TTTCTGAGACTGGGCGCTTTCTGTCCTGCGACGCGGCAACAACTACTGTGTGC 2322
Db 2421 TTTCTGAGACTGGGCGCTTTCTGTCCTGCGACGCGGCAACAACTACTGTGTGC 2480
QY 2323 ATGAGGACCATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2382
Db 2481 TTGAGGACCATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2540
QY 2383 GAGTACTTTGATCTCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2442
Db 2541 GATACTTTGATCTCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2600
QY 2443 AGGATGTCTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2502
Db 2601 AGGAGGCTCTTAAACCACTGACGCTGACGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2660
QY 2503 AAGCAGTGTAAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2562
Db 2661 AAGCAGTGTAAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2720
QY 2563 GAGCAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2622
Db 2721 GAACAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2780
QY 2623 TCACTGGGACACTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2682

Db 2781 TCATGGGACAGCTATGTTGGAGAGGTTGGGAGGCTAAGCGCCATAGAGAGGAGAACCTCC 2840
QY 2693 AACATAGAGGAGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722
Db 2841 AACACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880

RESULT 4
AK081643
LOCUS
DEFINITION
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130058L07 product: weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert sequence.
AK081643
VERSION
AK081643.1 GI:26349278
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
AUTHORS
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE
Functional annotation of a full-length mouse cDNA collection
NATURE 409 (6821), 685-690 (2001)
21085660
MEDLINE
11217851
REFERENCE
5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3734)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tgami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp].
 URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: <http://genome.gsc.riken.go.jp/>
 URL: <http://fantom.gsc.riken.go.jp/>

FEATURES Location/Qualifiers
 source
 1..3734
 organism="Mus musculus"
 mol_type="mRNA"
 strain="C57BL/6J"
 db_xref="PANTOM_DB:Cl30058107"
 db_xref="taxon:10090"
 clone="Cl30058107"
 tissue_type="head"
 clone_lib="RIKEN full-length enriched mouse cDNA library"
 dev_stages="16 days embryo"
 232..2859
 note="unnamed protein product; putative weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix] (SPR|Q90XB6, evidence: FASTY, 64.2%ID, 99.5%length, match=2601)"
 /codon_start=1
 /protein_id="BAC38279.1"
 /db_xref="GI:26349279"
 /translation="MAPPGLPLWLLSTALLSLLAGSAFLSHPLKGRFORDERNTRP
 NIILVTDQVEVSGMQVWVKTERIMEQGNHFINAFVTPMCCPSRSILTKGVH
 NHTYTNENSSSQAQHEKSRFPVFNSTGRFAFFKLINEINGSVTPPGKWE
 VGLKNRYNTLCRNGKEDSDYSLDILINDSVSPFRSKYKYPHPVIM
 VISAAHPGSDSAPQSRFLPNASQHTPSYNAFNPDKHWINRTGPKMTHMET
 NMLQRLQTLMSVDDSNETIYMLVETGELDNTYILYADHGVHIGQFLGVKGSMP
 YEFPIRVFVYRGVNPAGSLNPHI VLNIDLPILDILADLPAMDGKSLILKIDS
 EYRPNRFLHKKLWVDSFLVERGKLLHKRGDKVNAORENLPKYQVLDLCRAE
 YQTACEQLQXQVQVEDASGTLKHKCKGPMRGGGGSRALNLPKIDGQSEACS
 CDSGGGDYKUGLAGRKLFPKKYKTSYARNRSIRVAIEVDGEIYHVLDTVPQPN
 LSKPHWPGAPDQDDQSGFSFGTGLPDYSAPNPINVTYHRCVILENDVTQCDLDLYK
 SLOWKDKHLIHDEIETLQNKILNREVRGLKKRPERDCDCHRSYHSQKGRLLK
 KGSLSHPFRGLQKRLKLLREOKRKKLKLRLQNNDCRSMPLGTCFTHDNHHW
 QTAPLWTLGPPCACTSANNTYCLRTINETNLFCEPATGFIEYFDLSTDPYQLMN
 AVNTLDRDLNQLHVOIWELESCKGYQCNPRFNMDLGLRDGGSVEYQRFQRRKP
 EMKRPSSKSLQLWEHNEG"

CDS

BASE COUNT
 ORIGIN

956 a 1074 c 957 g 747 t

Query Match

62.2%; Score 2092; DB 11; Length 3734;

Best Local Similarity 86.1%; Pred. No. 0;
 Matches 2359; Conservative 0; Mismatches 360; Indels 21; Gaps 3;
 QY 1 GACTCCCGCATCCCAAAGAACCCAGATCAGCAAAAAGAAAGATGGCCCCCGGAGC 60
 Db 187 GGGCTCTGCATCAGGAAGAACGAAGGACCAAGCCACCCCAATGGACCCCTGGC 246
 QY 61 CTGGTCTCTGTCTGTGTCGCCAACTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
 Db 247 CTGCCACTATGGCTGTGTGTCCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 306
 QY 121 CTGTGCGCACACCGCTGCAAGCGAGGTTTCAGAGGCGCGGAGGAAATCCGCCCAAC 180
 Db 307 CTCTCCCATCCCCCGCTGAAGGAGCGCTTCCAGAGGCGCGGAGGAAATCCGCCCAAC 366
 QY 181 ATCATCTCTGTCTGTACGGGACACAGGATGTGGAGCTGGGTTCATGCAAGTGTATGAC 240
 Db 367 ATCATCTCTGTCTGTACGGGATGACAGGATGTGGAGCTGGGTTCATGCAAGTGTATGAC 426
 QY 241 AAGACCCCGCGCATCATGGAGCAGGGCGGGCGGACATTCATCAACGCCTTCTGTGACCA 300
 Db 427 AAGACAAGCGGTATCATGGAGCAGGGCGGGCGGACATTCATCAATGCTTCTGTGACTACA 486
 QY 301 CCAATGTGCTGTCCCGTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
 Db 487 CCAATGTGCTGTCCCGTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 546
 QY 361 ACTTACACCAACAATGAGAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
 Db 547 ACTTACACCAACAATGAGAAATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606
 QY 421 ACTTGTGCGCTGTACTCAATAGCACTGGCTACCGGACAGCTTTCTCTGGGAAGTATCTT 480
 Db 607 ACTTGTGCGCTGTATTTCAAAGCAAGCTTACCGGACAGCTTTCTCTGGGAAGTATCTT 666
 QY 481 AATGAATACACCGCTCTCTACGTCGCCACCGCTGGAAGGAGTGGGTCTCGGACTCTCTAAA 540
 Db 667 AATGAGTACAACCGCTCTCTACGTCGCCACCGCTGGAAGGAGTGGGTCTCGGACTCTCTAAG 726
 QY 541 AACTCGCGCTTTTAACTACACGCTGTGTGGAAACGGGTGAAAGAAAGACCGGCTCC 600
 Db 727 AACTCGCGCTTTTAACTACACACTCTGCGGAAATGGGGTGAAGGAAACATGCGCTCA 786
 QY 601 GACTACTCCAAGGATTACTCTACAGCTCTCATCAATATGACAGGCTGAGCTTTCTTTCGCG 660
 Db 787 GACTACTCCAAGGATTACTCTACAGCTCTCATCAATATGACAGGCTGAGCTTTCTTCCGA 846
 QY 661 AGTCCAAAGATGTACCGGACAGGCGAGTCTCTATGTTGTCATGAGCATGAGCCGCTCC 720
 Db 847 ACATCCAAGAGATGTACCCACACAGGCGCGGTCTCTATGTTGTCATGAGCATGAGCCGCTCC 906
 QY 721 CACGCGCTGTAGGATTCAGCGCCCAATATTACAGCTCTTCCCAACGCTATCTCAGCAC 780
 Db 907 CATGCGCGCGGAGTCTAGCAGCCCGCTCTCTACGGCTCTTCCCAATGCTGCCAGCAC 966
 QY 781 ATCAAGCGGAGTACAATCTACGCGCCCAACCGGCAACAACTGGATCATCGCTTACACG 840
 Db 967 ATCAAGCGGAGTACAATCTACGCGCCCAACCGGCAACAACTGGATCATCGCTTACACG 1026
 QY 841 GGGCCCATGAGCCCATCCATGGAATTCACCAACATGCTCCAGCGGAGGCGCTTTCAG 900
 Db 1027 GACCCCATGAGCCCATTCATGGAATTCACCAACATGCTCCAGCGGAGGCGCTTTCAG 1086
 QY 901 ACCCTCATGTCTGTGAGAGTCTCCATGAGACGATTTACAACTGCTGTTTGAACGCGGC 960
 Db 1087 ACCCTCATGTCTGTGAGAGTCTCCATGAGACGATCTATGACATGCTGGTGGAGACGCGG 1146
 QY 961 GAGCTGGACAACAGTACATCGTATACACCGCGGACACCGTTCACCATCGGCGAGTTT 1020
 Db 1147 GAGCTGGACAACAGTACATCGTATACACCGCGGACACCGTTCACCATCGGCGAGTTT 1206
 QY 1021 GGCCTGGTGAAGGGAATCCATGCGCATATGAGTTTGACATCAGGGTCCCGCTTCTACGTG 1080

Db 1207 GGCTGGTGAAGGCGAAGTCTATGCGGTATGAATTCGACATCAGAGTCCGCTTCTACGTG 1266
QY 1081 AGGGGCCCAACGCTGAAGCGGCTCTCTGAATCCCAATCGTCTCCTCAACATTCGACCTG 1140
Db 1267 AGGGGCCCAACGCTGAAGCTGGCTCTCTGAACCCCACTTGTCTCAACATTCGACCTG 1326
QY 1141 GCCCCACCAATCCTGACATTTGAGGCTTGGACATACCTTGGGATATGAGCGGGAATCC 1200
Db 1327 GCCCCACCAATCCTGATATCGTGGACTTGGACATCCCTTGCAGACATGAGCGGGAATCT 1386
QY 1201 ATCCTCAAGCTGCTGACACGAGCGGCGGTGAATCGGTTTCACTTGAAGAAGAGATG 1260
Db 1387 ATTCTCAACTACTGACTCAGAGCGGCGAGTGAACCGTTCCACTTGAAGAAGAGCTG 1446
QY 1261 AGGGTGTGGCGGACTCTCTTCTTGTGGAGAGGCAAGCTCTACACAGAGAGACAAAT 1320
Db 1447 AGGGTGTGGCGGACTCTCTCTCTGTGGAGAGGCAAGCTCTCTCCACAGAGGAGGGT 1506
QY 1321 GACAGGTGGACGCCACGAGGAGAGACTTCTGCGCCNAGTACAGCGTGTGAGGACCTG 1380
Db 1507 GACAAAGTGAATGCCCAGGAGAGAACTTCTCTGCGCAAGTACACGCGGTGAGGACCTG 1566
QY 1381 TGTACGCTGCTGAGTACCAGACGCGGTGTGAGCAGCTGGGACAGAAGTGGCAGTGTGTG 1440
Db 1567 TGTACGCGAGCTGAGTACCAGACAGCATCGAAGCAGCTGGGCGAGAGTGGCAGTGTGTG 1626
QY 1441 GAGGACGCCACGGGGAAGCTGAAGCTGCAATAGTGCAGAGGCGCCCATCGGCT----- 1493
Db 1627 GAGGACGCTTCTGGGACGCTGAAGCTGCAAAATGTAAAGGCGCCCATCGGCTTTGGTGGC 1686
QY 1494 --GGGCGGACAGAGCCCTCTCCAACTCGTGCACAGTACTAGCGGCGAGGCGAGCGAG 1551
Db 1687 GCGGTGGCAGAGAGCCCTCTCCAACTGTGTGTCGCGCCNAGTATGAGCGGCGAGGCGAG 1746
QY 1552 GCTGCACTGTGACA-----GCGGGGACTACAAGCTCAGCTGCGCGGACGCGCGG 1602
Db 1747 GCTGCACTGTGACAGTGGCGGTGGAGGGGACTACAACTGGGCTGGCTGGACGCGG- 1805
QY 1603 ABAARACTCTCAAGAGAGTACAGGCGCCGCTATGTCGCGAGTGCCTCCATCGCTCA 1662
Db 1806 --TAACTCTTTAAGAAAAGTATAAGACAGCTATGCCCGGAACCGCTTCCATCGTTCC 1863
QY 1663 GTGGCCATCGAGTGGACGCGAGGCTGTACACGCTAGGCTGGGTGATGCGCGCCAGCGCC 1722
Db 1864 GTGGCCATCGAGTGGACGCTGAGATATACACGCTAGGCTTGGATCTGTGCTCAGGCC 1923
QY 1723 CGAACTCACAGGCGGCACTGGCCAGGCGGCTTGAAGCCAGAGTACAGGATGGT 1782
Db 1924 CGCACTTAGCAAGCGGCACTGGCCAGGCGGCTTGAAGCCAGAGTACAGGATGGT 1983
QY 1783 GGGGACTTCAGTGGCACTGGAGGCTTCCCGACTACTCAGCGCGCCAAACCCCATTAAGTG 1842
Db 1984 GGCAGTTTCAGTGGTACTGTGGCTTCCAGATTATCTGCGCCCAATCCCATCAAGTG 2043
QY 1843 ACACATGGTGTACATCTTAGAAGACGACAGCTCAGTGTGACTGACCTGTACAG 1902
Db 2044 ACCATCGGTGTACATCTTGAAGATGACAGCTCAGTGTGAGCTGACCTGTACAG 2103
QY 1903 TCCTCGAGGCTTGAAGACACACAGCTGCATCTGACCAAGATTGAACCTGTGAG 1962
Db 2104 TCCTCGAGGCTTGAAGACACACAGCTGCATCTGACCATGAGATGAAACCTGTGAG 2163
QY 1963 AACAAATTAAGAACTGTAGGGAAGTCCGAGTCACTTGAAGAAAAGCGGCCAGAGAA 2022
Db 2164 AACAAATTAAGAACTTCCGAGAGTCCGAGGCTCACCTGAAGAAAGCGACCGGAGAA 2223
QY 2023 TGTGACTGTACAAATACAGTACCAACCGGACAGGCGGCTTCAAGCAGAGGCG 2082
Db 2224 TGTGACTGCCATAGAAATCAGTTACCAACGCGCAACCAAAAGGCGGTCTCAAGCAGAGGCG 2283
QY 2083 TCCAGTCTGCATCTCTTCCAGGAGGCGCTTCCAGAGAGAGGACAAAGTGTGGCTGTGGG 2142
Db 2284 TCCAGCTTGACCCCTTTCAGGAGGCTCTGCGAGGAGAGGACAAAGTGTGGCTGTGGG 2343

QY 2143 GAGCAGAGCGCAAGAGAAACTCCGCAAGCTGCTCAAGCGCTGCAGAACCAACGACAG 2202
Db 2344 GAGCAGAAACGCAAGAGAAACTGCGCAAGCTGCTCAAGCGCTGCAGAACCAACGATAG 2403
QY 2203 TGAGCATGCCAGGCTCAGCTGCTTACCCAGCAACACAGCACTGGCAGACGGCGCT 2262
Db 2404 TGAGCATGCCGCGCTCAGCTGCTTACCCAGCAACACCACTGGCAGACGGCGCA 2463
QY 2263 TTCTGGACACTGGGCGCTTCTGCTGCTCACAGCGCCAAACAATAACAGTACTGCTGC 2322
Db 2464 CTCTGAGCGTGGGCGCTTCTGCGCTGACAGCGCCAAACAACAACAGTACTGCTGC 2523
QY 2323 ATGAGCAACATCAATGAGACTCAAAATTTCTCTTCTGTAATTTGCAATCGCTTCTCTA 2382
Db 2524 TTGAGGACCAATAATGAGACCCCAAACTTCTCTCTGCGAATTTGCAACCGCTTCTATA 2583
QY 2383 GAGTACTTTGATCTCAACAGACCCCTTACAGCTGATGAATGCAATGCAACACACTGGAC 2442
Db 2584 GAATACTTTGACCTCTAGTACAGACCCCTTACAGCTGATGAACCGGTGAACACACTGGAC 2643
QY 2443 AGGATGTCTCTCAACAGCTTACACGTACAGCTCATGAGCTGAGGAGCTSCAAGGTTAC 2502
Db 2644 AGGAGCTCTTAAACCACTGCACTGCACTGAGCTGAGGAGCTGTAAGGCTAC 2703
QY 2503 AAGAGTGTAAACCCCGGACTCGAAACATGAGCTGGGACTTAAAGATGAGGAGACTAT 2562
Db 2704 AAGCAGTGCAAACCCCGGACCCGCAACATGAGCTTGGGCTTTAGAGACGGAGAGCTAT 2763
QY 2563 GAGCAATACAGGAGCTTTCAGCGTCCAAAGTGCCAGAAATGAAGAGACTTCTTCCAAA 2622
Db 2764 GAACATACAGGAGCTTTCAGCGTCCAAATGCCAGAAATGAAGAGACTTCTTCCAAA 2823
QY 2623 TCCTGAGGACAACTGTGGAGAGCTGGGAGGTAAAGAAACAACAGAGTGGACCTCCAA 2682
Db 2824 TCCTGAGGACAGCTATGGAAGGCTGGGAGGCTAAGCGCCATAGAGAGAGAACTCC 2883
QY 2683 AAACATAGAGGACATCACTGACTGACGACGAGCAATGAATA 2722
Db 2884 AAACAGGCGGCTCTGTGTGCTGCCAGGCCATGCAAAA 2923

RESULT 5
AK049170
LOCUS
DEFINITION
Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone: C30011A01 product: weakly similar to
N-ACETYLGUCOSAMINE-6-SULFATASE [Coturnix coturnix], full insert
sequence.
AK049170
VERSION
AK049170.1 GI:26339841
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM

REFERENCE
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, K., Okazaki, Y., Gojibori, T., Bono, H., Kasakawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Saito, R., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kaniya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S., and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2685)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 2685
/organism="Mus musculus"
/mol_type="mRNA"

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

FEATURES
source

```

/strain="C57BL/6J"
/db_xref="PANTOM DB:C330011A01"
/db_xref="taxon:10090"
/clone="C330011A01"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
<1.1816
/note="unnamed protein product; putative
weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE
[Coturnix coturnix] (SPTR|Q90X86, evidence: PASTY,
64.2%id, 99.5%length, match=2601)"
/codon_start=2
/protein_id="BAC33584.1"
/db_xref="GI:26339842"
/translat="HMEFTNMLQRKRLTQMSVDSMETIYDMLVETGELDNTLYLT
ADHGVIQGFGLVKGSMPEYDFIRVPYVGRGNVVEAGSLNPHIVLNDLAPTILIDIA
GLDIPADMDKSLKLLDSEPVNFEHLKKLVWRDSFLVGRKLLHKGSGDKVNAQ
BENFLPKYQVXDLCQRAEYQTACQLQKQWCVDEASGTLKHCKGKPMFGGGGS
RALSNLVPIYDQSSACSDSGGGDYGLAGRRKLFKKYKTSYARNSSIRSAI
EVDGEIYHVLDTVPQPNLSKPHWGPEDDDKDGSGFTGGLPYSPNFIKVT
HRCYILENDTVQCDLDLYSLQAWKHKLHIDHEITLQNKIKLREVRGLKKRPE
BCDCHRSYHSGHRLKHGSLHPPFKGLQEKDKVLLREQKRLKRLKRLQON
NDTCSMPLGTCFTDHNHMTAPLWTLGPPFCACTSANNNTYCLRTNETNFIPLCFE
ATGPTLEVLDLSTDYQLMNAVNTLDRDVLNQLHVLMLRSCGKYKQCNPRTRNDLG
LRDGSYEQYRQPRKRWPMKRFSSKSLGLWEGWEG"
BASE COUNT      710 a  727 c  700 g  548 t
ORIGIN
Query Match      42.5%; Score 1430.8; DB 11; Length 2685;
Best Local Similarity 86.3%; Pred. No. 1.5e-285;
Matches 1624; Conservative 0; Mismatches 237; Indels 21; Gaps 3;
QY      859  CACATGGAAATTCACCAACATGCTCCAGCGGAGCGCTTGCAGACCTCATGTCGGTGGAC 918
Db      2   CACATGGAAATTCACCAACATGCTCAACGAAAGCGCTTACAGACCTCATGTCGGTGGAT 61
QY      919  GACTCCATGGAGACGATTTACAACATGCTGCTGCAGCGGCGAGCTGGACAACACGCTAC 978
Db      62  GACTCCATGGAGACGATTTACATGCTGCTGCAGCGGCGAGCTGGACAACACGCTAC 121
QY      979  ATCGTATACCGCGGACCGGTTTACACATCGCCGCGGAGTTTGGCTTGGTGAAGGGA 1038
Db      122  ATCTGTACACCGCGGACCGGTTTACACATCGCCGCGGAGTTTGGCTTGGTGAAGGGAAG 181
QY      1039  TCCATGCCATATGATTTGACATCAGGCTCCGCTTCTACGTGAGGGGCCCCCAACGCTGAA 1098
Db      182  TCTATGCGGTATGAATTCGACATCAGAGTCCCGTCTTACGTGAGGGGCCCCCAACGCTGAA 241
QY      1099  GCGGCTCTCTGAATCCCAACATCGTCTCAACATTCACCTGGGCCCCCACCACCTCTGGAC 1158
Db      242  GCTGGCTCTCTGAACCCCAACATTCGCTCAACATTCGCTGGGCCCCCACCACCTCTGGAT 301
QY      1159  ATTGAGCGCTGGACATACCTCGGATATGACGCGGAAATCCATCTCAAGCTCTGGAC 1218
Db      302  ATCGCTGAGCTGGACATCCCTGCGACATGCGGGAAGTCTATTCTCAAACTACTGGAC 361
QY      1219  ACGGAGCGCGGCTGAATCGGTTTTCATTGAAAAGAGATGAGGGTCTGGCGGACTCC 1278
Db      362  TCAGAGCGGCGAGTGAAACCGGTTCCATTGAAAAGAGCTGAGGGTCTGGCGGACTCC 421
QY      1279  TTCTTGGTGGAGAGGCAAGCTCTCTACACAAGAGACAATGACAAGGTGGAGCGCCAG 1338
Db      422  TTCTTGGTGGAGAGGCAAGCTCTCTCACAAGAGGAGGCTGACAAAGTGAATGCTCCAG 481
QY      1339  GAGGAGACTTTCTGCCCAAGTACAGCGGTGTGAGGACCTGTGTGACGCTGTGAGTAC 1398
Db      482  GAGGAGAACTTCTTGGCCCAAGTACAGCGGTGTGAGGACCTGTGTGACGAGCTGTGATAC 541
QY      1399  CAGAGCGGCTGTGAGAGCTGGGACAGAGTGTGTGTGGAGGACGCGCACCGGGAAG 1458
Db      542  CAGAGCATCGGAACAGCTGGGCGAGAGTGTGTGTGGAGGACGCTTCTGGGACG 601
QY      1459  CTGAAGCTGCATAGTGCAGAGGCGCCCATCGGGCT-----GGCGGCGAGCAGAGCC 1509

```

Db 602 CTGAAGCTGCACAAATGTAAGAGCCCAATGCGTTTGGTGGCGGTGGCAGCAGAGCC 661
Qy 1510 CTTCTCCAACTCGTCCCAAGTACTACGGCAGGCGAGGCGCTGCACTGTGACA-- 1567
Db 662 CTCTCAACTCTGTGCTCCCAAGTATACGGCCAGAGCAGGAGGCGCTGAGCTGTGACGT 721
Qy 1568 -----GGGGGACTACAACTCAGCTGCGCGGAGCGCCGAAAGAACTCTTCAAGAAG 1620
Db 722 GCGGTGGAGGGGACTACAACTGCGGCTGCTGCGAGCGCG---TAAGCTCTTTAAGAAA 778
Qy 1621 AGTACAGGCGCAGTATGTCGCGAGTGCCTCCATCCCTCAGTGGCCATCAGGTGGAC 1680
Db 779 AAGTAAAGCAGCTATGCTCCGGAACCGCTCCATCCGTTCCGTGGCCATCAGGTGGAC 838
Qy 1681 GGCAGGGTGTACAGGTAGGCGCTGGGTGATGCGGCCCGAGCCCGGAAACCTCACCAAGCGG 1740
Db 839 GTGAGATATACCACTAGGCTTGGATCTGTGCTCAGCCCGCACTTACCAAGCGG 898
Qy 1741 CACTGCCAGGCGGCGCTGAGGACCAAGATGACAGGATGTTGGGACTTCAGTGGCACT 1800
Db 899 CACTGCCAGGCGGCGCTGAGGACCAAGATGACAGGATGTTGGGACTTCAGTGGTACT 958
Qy 1801 GGAGGCTTCCGACTACTCAGCGCCCAACCCATTAAAGTGACACATCGTGTACATC 1860
Db 959 GTGGCTTCCAGATATTTCTGCCCCCAATCCATCAAGTGACCCATCGTGTACATC 1018
Qy 1861 CTAGAGACGACACAGTCCAGTGTGACTGACCTGTGTACAGTCCCTGAGCGCTGGAAA 1920
Db 1019 CTTGAGATGACACAGTCCAGTGGACTTGGACCTGTACAAAGTCCCTGAGCGCTTGGAAA 1078
Qy 1921 GACCAAGCTGCATCGACACCAAGATGAAACCTTGCAGAACAAATTAAGAACCTG 1980
Db 1079 GACCAAGCTGCATCGACACCAAGATGAAACCTTGCAGAACAAATTAAGAACCTT 1138
Qy 1981 AGGGAAGTCCGAGTCACTGTAAGAAAAGCGCGCCAGAGGAATGTGACTGTACAAAATC 2040
Db 1139 CGAGAAGTCCAGGGTCACTGTAAGAAAGAGCGACCGGAAGATGTGACTGCCATAGAATC 1198
Qy 2041 AGCTACCAACCCAGCACAAGGCGCTTCAAGCAGACAGGCTCCAGTCTGCATCCCTTC 2100
Db 1199 AGTTACCAACCCAGCACAAGGCGCTTCAAGCAGCACAAGGCTCCAGCTGCACCCCTTC 1258
Qy 2101 AGGAAGGCGCTGCAAGAGAGCAGCAAGTGTGGCTGTGGCGGAGCAGAGCGAAGAG 2160
Db 1259 AGGAAGGCTGCGAGGAGAGCAGTGTGGCTGTGGCGGAGCAGAGCGAAGAG 1318
Qy 2161 AAATCCGCAAGTGTGCTGAGGCTGCGAGAACAGACAGTGTGAGCATGCGAGGCTC 2220
Db 1319 AAATCCGCAAGTGTGCTGAGGCTGCGAGAACAGACAGTGTGAGCATGCGAGGCTC 1378
Qy 2221 AGTGCTTCCACCAAGCAGACAGTGTGGCGGCGCTTCTGACACTGCGGCGCT 2280
Db 1379 AGTGCTTCCACCAAGCAGACAGTGTGGCGGCGCTTCTGACACTGCGGCGCG 1438
Qy 2281 TTCTGTGCTGCTGACAGGCGCCCAATACAGTGTGCTGCTGAGCAGCAGCAGCAGT 2340
Db 1439 TTCTGTGCTGCTGACAGGCGCCCAATACAGTGTGCTGCTGAGCAGCAGCAGT 1498
Qy 2341 ACTCAATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 1499 ACCCAATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1558
Qy 2401 ACAGAGCCCTTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 1559 ACAGAGCCCTTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1618
Qy 2461 CTACAGTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 1619 CTGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1678
Qy 2521 ACTCGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580

Db 1679 ACCCGCAACATGCACTGGGCTTAGAGACGAGGAGGAGCTATGAACAATACAGGCGAGTTT 1738
Qy 2581 CAGCGTCGAAAGTGGCCAGAAATGAGAGACCTTCTTCAATCTACTGGGCAACTGTGG 2640
Db 1739 CAGCGTCGAAATGAGGAGAAATGAGAGACCTTCTTCAATCTACTGGGCAAGCTATGG 1798
Qy 2641 GAAGGCTGGGAAGGTTAAGAAACAACAGAGGTGGACCTTCAAAAAACATAGAGGCACTACC 2700
Db 1799 GAAGGTTGGGAAGGTTAAGCGCCATAGAGAGAGGAACTCTCAAAACACGAGGCGCTCGT 1858
Qy 2701 TGACTGCACAGGCAATGAAAA 2722
Db 1859 TGCTGCCAGGCGCATGCAAAA 1880

RESULT 6
AK045002 4564 bp mRNA linear HTC 05-DEC-2002
LOCUS
DEFINITION
Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length
enriched library, clone:B130018L10 product:SULFATASE FP homolog
[Rattus norvegicus], full insert sequence.
AK045002
ACCESSION
VERSION AK045002.1 GI:26336990
KEYWORDS
SOURCE HTC; CAP trapper.
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS
TITLE
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
MEDLINE
PUBMED 99279253
PUBMED 10349636
2
REFERENCE
AUTHORS
TITLE
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
MEDLINE
PUBMED 20499374
PUBMED 11042159
3
REFERENCE
AUTHORS
TITLE
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Warahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
MEDLINE
PUBMED 20530913
PUBMED 11076861
4
REFERENCE
AUTHORS
TITLE
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, K., Harada, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Arakawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,

Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohrsuiki, S., and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 MEDLINE
 PUBMED
 11217851
 5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 4564)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tonari, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. Tomohiro Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Funako Atsugi City, Kanagawa Prefecture, Japan) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1. 4564
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:B130018L10"
 /db_xref="taxon.10090"
 /clone="B130018L10"
 /tissue_type="parthenogenote"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="9.5 days embryo"
 382. 2970
 /note="unnamed protein product; SULFATASE FP homolog [Rattus norvegicus] (SPTF) AAL71906, evidence: PASTY, 96.4%id, 100%length, match=2610)
 putative"

/codon_start=1

/protein_id="BAC32179.1"

/db_xref="GI:26336991"

/translation="MKYSLWALLAVLGTQLLSGLSTVRSORFRGRLOERKNIRPN
 IILVLDODVEIGSLQWNKTKIMEQGNFTNAPVTPMCCPSRSRSMGLTKVHN
 HNYVTNNSSSSWQMEHPRFTAVILANTGRTAFPGKILNENSTIPFGWEHL
 GLKNRFPYVCRNGEKHGFDAKDYFDLITNESINYPFKSRKMPHPHPIMV
 ISHAHPGSDAPQSKLYPNASOHTTFSYANPNMDKIMWQTMPLTHMEFTN
 VLQRKQLTMSVDSIPGVLYNNMVSGBLDNTYIITADHGVHIGQFLVKGSMPT
 DFQIRVPPTRGSPISPGSVPOIVLNIDLPILDIAGLSDPSDVGSLKLDLE
 KPGNRRTNKKAKIMRDTLVERKGLRKESGKNIQSNHLPKRVKELCOQARY
 QTACEPGQNWQIEDTSGKLRIHKCKGSPSLLTVRQNRNLYSRGLHDKDKRCHRD

SGVRSRSORKNORQFLRNKGTPKYKPRFVHTROTSLRSVEFECEIYDINLEEBELQV
 LPPSLAKHDEGHGGTGHQAAGDINENMLADNNAVGLPATRVTHKCPILPNT
 IHCRRELYQSARAWKHDKAYIDKEIVLQDKLKNLREVRGHLKKRKEEGCQSQSY
 NKEGKVRQEKLSHLHPFKAAQEVDSQKLQKFKHRRRKRKRKRQKRGCECSL
 PGLCTFTDHHNQWTFAPFNWLSFCACSTNNNTYWCILRTVNTNHLFCFETATGLE
 YFMTNTDPIYQLTNTVHTVERSILNQLHIQLMELSCQGYKQCNPRKSLDTGAKEGN
 YDPHRA"

BASE COUNT 1282 a 1079 c 1150 g 1053 t
 ORIGIN

Query Match 31.6%; Score 1062.4; DB 11; Length 4564;
 Best Local Similarity 65.7%; Pred. No. 2.3e-209;
 Matches 1666; Conservative 0; Mismatches 811; Indels 57; Gaps 6;

QY 97 CTGCTGGGTGGAAGCTCGGCCTTCTGTCGCGACACCGCCTGAAAGGAGGTTTCAGAGG 156
 DB 430 CTGCTGGGAAGCCTGTGTTCACCGTTCGCTCCAGAGGTTCCGAGGAGATACAGCAG 489
 QY 157 GACCGCAGGAATACCGCCCAACATCATCTGTGTGTGAGCGGACGACGAGATGTGGAG 216
 DB 490 GAACGAAAAACATCAGACCCCAATATTATCTTGTGTCTACTGACGACGAGATGTGGAG 549
 QY 217 CTGSGTTCATGCGAGGTGATGAACAAGACCGCGCATCATGTGAGCGAGCGGGGGCGGAC 276
 DB 550 CTGSGTTCCTGCAAGTCAATGAACAAGAGATCATGGAACAGGGTGGGGCCACC 609
 QY 277 TTCATCAAGCGCTTCGTGACACACCCATGTGTGTCCTCCCTCAAGCTCTCTCCATCTCACT 336
 DB 610 TTCACCAATGCTTTGTGACACCGCCCATGTGTGTCTCATCAGCTCATCATCTCACT 669
 QY 337 GGCAGTACGTCACACCAACCAACACCTACACCAACCAATGAGAACTGCTCTCGCCCTCC 396
 DB 670 GGGAGTAGTACGTGCATAACCAATGTCTACACCAACCAATGAGAACTGCTCTCGCCCTCC 729
 QY 397 TGGCAGGCACACGACGAGCGGCACTTTGCGGTGTACCTCAATAGCACTTGGCTACCGG 456
 DB 730 TGGCAGGCAATGACGAGCGCTCGACCTCTGCTGTATCTCAACAACACCGGCTACAGA 789
 QY 457 ACAGTTTCTTCGGGAGATATCTTAATGATACAAACGCGTCTTACGTCGCAACCGGCTGG 516
 DB 790 ACAGCTTTTGTGAAAATACCTCAATGAATACAAATGCGAGCTACATCCCTCTCGGATGG 849
 QY 517 AAGGAGTGGGTGCGACTCTTTAAAACTCCCGCTTTTATAACTACACGCTGTGTGGGAC 576
 DB 850 CGAGATGGCTCGGATTAATCAAGAACTCTGTTCTATTAATACATCTGTTTGTGGCAAC 909
 QY 577 GGGGTGAAGAAAAGCAGCGCTCCGACTACTTCAAGGATTACTTCAAGACCTCATCAACC 636
 DB 910 GGCATCAAGGAGAAGCATGGATTGATTATGCAAGGATTACTTCAAGACTTAATCACT 969
 QY 637 AATGACGCTGAGCTTCTTCGCGACGTCGAGAGATGTACCGGACAGCGGACGAGCTCCTC 696
 DB 970 AACCAGAGCATTAATTAATCTTCAAAATGTCTAAGAGAAATGATCCCATAGGCCCATATG 1029
 QY 697 ATGTCATACGCCATCAGCGCCCGCCCGCGCTGAGGATTTCAGCGCCCACTAATATTCACGC 756
 DB 1030 ATGTCATCAGCCATCAGCAGCACCACATGGCCCTGAAGACTCCGCCCCACAGTTTCAAAA 1089
 QY 757 CTCTTCCCAAGGATCTCAGCAGCATCAGCGGAGCTTACAACTACGCGCCCAACCCCGGAC 816
 DB 1090 CTGTACCCCAATGCTCTCCCAACACATAACCCAGCTATAACTATATGCAACCAATGAGAC 1149
 QY 817 AACACATGGATCATGCGCTACACGGGCGCCCATGAAGCCCATCAACATGGAATTCACCAAC 876
 DB 1150 AAACTGGATCATGAGTACACAGGGCCCATGTTGCTATCATCATGGATTCACCAAC 1209
 QY 877 ATGCTCAGCGGAAGCGCTTGCAGACCCCTCATGTGCGTGGAGCACTCATGAGGAGCAATT 936
 DB 1210 GTCTCCAGCGCAACCGCAGACTCTCATGTGCTGCTGACGACTCTGTGCGAGGTTG 1269
 QY 937 TACAACATGCTGGTTGAGCGGGGAGCTGGAACAACACGCTACATCGTATACACCGCGGAC 996
 DB 1270 TACAACATGCTTGTGGAGAGCGGGGAGCTGGAACAACACTTACATCATCTACATCGCGGAC 1329

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

20530913
 11076861
 4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuenli, F., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schrim, L.M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.P., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyoko-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Haegawa, Y., Kawai, H., Kohseki, S. and Hayashizaki, Y.

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

21085660
 11217851
 5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

6 (bases 1 to 4613)
 Nature 420, 563-573 (2002)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.riken.go.jp/
 URL: http://fantom.riken.go.jp/

FEATURES
 source

Location/Qualifiers

1..4613

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:8030446E18"

/db_xref="taxon:10090"
 /clone="8030446E18"
 /sex="male"
 /tissue_type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="15 days embryo"
 misc_feature
 1..4613
 /note="SULFATASE FP homolog [Rattus norvegicus]
 (SPR|AAL71906, evidence: FASTY, 96.4%ID, 100%length, match=2610)"
 BASE COUNT 1321 a 1083 c 1136 g 1073 t
 ORIGIN

Query Match 31.6%; Score 1062.4; DB 11; Length 4613;
 Best Local Similarity 65.7%; Pred. NO. 2.3e-209;
 Matches 1666; Conservative 0; Mismatches 811; Indels 57; Gaps 6;

QY 97 CTGCTGGGTGGGAAGCTCGGCCCTTCCTGTCGACCAACCCGCTGAAAGGAGGTTTCAGAGG 156
 Db 520 CTGCTGGGAAGCTGTGTTCACCGTTCGGTCCAGAGGTTCCGAGGAAGATACAGCAG 579
 QY 157 GACCGAGGAACATCCGCCCAACATCATCTGTGTGTCGACGACGACGACGATGGAG 216
 Db 580 GAACGAAAAAACATCAGACCCCAATATTATCTGTGTCTCACTGACGACGACGATGGAG 639
 QY 217 CTGGGTTCCATGCAAGTGTATGACGACGACGACGACGACGACGACGACGACGACGAC 276
 Db 640 CTGGGTTCCCTGCAAGTGTATGACGACGACGACGACGACGACGACGACGACGACGAC 699
 QY 277 TTCTCAACAGCCCTTCGTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 336
 Db 700 TTCAACCAATGCTTTGTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 759
 QY 337 GGCAGTAGCTGTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 396
 Db 760 GGCAGTAGCTGTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 819
 QY 397 TGGCAGGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 456
 Db 820 TGGCAGGACCAATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 879
 QY 457 ACAGCTTTCTTGGGAAGTATCTTAATGAATACAAAGGCTCTTACGTCACACCGGCTGG 516
 Db 880 ACAGCTTTCTTGGGAAGTATCTTAATGAATACAAAGGCTCTTACGTCACACCGGCTGG 939
 QY 517 AAGGAGTGGTGGGACTCTCTTAAAGACTCCCGCTTTTAACTACAGCTGTGCGGAC 576
 Db 940 CGAGTAATGGCTGGGATTAATCAAGAAATCTCTGTTCTAATTAATTAATTAATTAAT 999
 QY 577 GGGGTGAAGAAAAGCAAGCGGCTCCGACTACTTCCAAAGGATTAACCTCACAGACCTCATACC 636
 Db 1000 GGCATCAAGGAGAGCATGGATTTGATTATGCAAGGATTAATTTACAGACTTAATCACT 1059
 QY 637 AATGACGGGTGAGCTTTCTCCGACGCTCCAGAGAGTGTACCCGACAGCGCAGTCTC 696
 Db 1060 AACGAGAGCATTAATTAATTTCAAAATGTCTAAGAGAAATGTATCCCATAGGCCCATATG 1119
 QY 697 ATGTGTCATCAGCCATGACAGCGGCTCCAGGCTTCAGGATTCAGGCCCCCAACAATATTCACGC 756
 Db 1120 ATGTGTCATCAGCCATGACAGCCATGACGACGACGACGACGACGACGACGACGACGAC 1179
 QY 757 CTCTTCCCAAAAGCATCTCAGCAGCATCAGCGGAGGTACAACTAGCGGCCCAACCGGAC 816
 Db 1180 CTGTACCCCAATGCTCTCCCAACACATAAACCCAGCTATAACTATGACCCGAAACATGAC 1239
 QY 817 AAACACTGGATCATGCGGTACACGCGGCCCCATGAAGCCCATCCACATGGAATTCACCAAC 876
 Db 1240 AAACACTGGATCATGCGGTACACGCGGCCCCATGTTGCTATCCACATGGAGTTTCCACAC 1299
 QY 877 ATGTCTCAGCGGAAGCGCTTCAGACCGCTCATATGTCGGTGGAGCACTCCATGGAGACGATT 936
 Db 1300 GTCTCTCCAGCGCAACGACTCGAGACTCTGATGTCAGTTGACGAGCACTCTGTCGAGAGGTTG 1359

QY 937 TACACATCTGCTGGTTGAGACGGCGAGCTGGACACACGATCATCGTATATACACGCCGAC 996
Db 1360 TACAAATCTGTTGGAGACGGCGAGCTGGACACACCTTATCATCTATACACTGCGGAC 1419
QY 997 CACGGTTACACATCGGCTGAGTTGGCTGTTGAAAGGGAATCCATCCATATGAGTTT 1056
Db 1420 CACGGCTACACATCGGCTGAGTTGGCTGTTGAAAGGGAATCCATCCATATGAGTTT 1479
QY 1057 GACATCAGGGTCCGTTCTACGTTGAGGGGCCCAACGTTGGAAGCCGGTGTGTGAATCCC 1116
Db 1480 GACATCCGTTGCTCTTCTTCTATCTGCTGTCACCAAGATAGAACCAAGGGTTCGATAGTCCCA 1539
QY 1117 CACATCGTCTCAACATTCAGCTGGCCCCCACCATCTCGGATTCGAGCTGCGGGCTGGACATA 1176
Db 1540 CAGATTGTTCTGAACTACGACCTAGCTCTCTCATCTCGATATTCGTTGCTTGAATCT 1599
QY 1177 CTTGCGGATATGGAACGGGAAATCCATCTCAAGCTGCTGGACACGGAGCGGCGGTGAAT 1236
Db 1600 CTTCTGATGTGATGCTAAATCTGCTCTCAAACTTCTAGACCTCGGAAAGCCGGTTAAC 1659
QY 1237 CGGTTTCACTTGAAGAGAGATGAGGTCTGGCGGACTCTCTTCTGTTGGAGAGAGGC 1296
Db 1660 AGGTTTGGACAAACAAAGAGGCAAAATCTGGCGTGATACATTCCTAGTGAAGAGAGG 1719
QY 1297 AGCTCTCTACAAAGAGAGACATGACAAAGGTGGACGCCCGAGGAGGAACTTTCTGCC 1356
Db 1720 AAATTTTACGGAAGAGAGGATCCGSCAAGATATCCACAGTCAACCACTTGCC 1779
QY 1357 AGTACACAGCTGTGAAGACCTGTGTAGCTGTGAGTACAGACGGGTGTGAGCAG 1416
Db 1780 AAATATGAGAGGCTCAAGGAACTGTGCCAGAACGACATACACAGACGCTGTGAGCAG 1839
QY 1417 CTGGGACAGAAGTGGCAGTGTGGAGGACGCCACGGGGAAGCTGAAGCTGATAGTGC 1476
Db 1840 CTTGGGCAACATGGCAGTGTCATCGAGGACATCGGCAAGCTCCGAATCCAAAGTGT 1899
QY 1477 AAGGCCCCA-----TGCGGTGGCGGCGCAGACAGAGCCCTTCCAACTGTGCCCCAAG 1530
Db 1900 AAGGGAACCCAGGACCTGTCTACGGTCCCTCAGAATGCACGCAACCTCTACTCTCGCGGA 1959
QY 1531 TACTACGGCAGGACGAGCGCTGCACCTGTGACAGGGGACTACAGCTAGCCTG 1590
Db 1960 TTGCATGACAAAGACAAAGAGTGCCATTGTAGGACTGTGGCTATCGTCCAGCAGAAAGC 2019
QY 1591 GCCGACGCCGGAATAAATCTTTCAAGAG-----AAGTACAAAGCCACG 1635
Db 2020 CAGAGAAAGAAATCAAAGGCAGTTCTGAGGAAACAAAGGAAACACCAAGATATAAGCCAGA 2079
QY 1636 TATGTCGCGCAGTCCGCTCCATCCGCTCAGTGGCCATCGAGGTGACGGCAGGGTGTACCAC 1695
Db 2080 TTTGTGCACACCCGCGAGACTCGGTCCTCTGTGTGAATTCGAAGGTGAATATATGAC 2139
QY 1696 GTAGGCTCGGTGATGCCGCCAG-----CCCCGAAACCTCAACCAAGGGGAC 1743
Db 2140 ATCAACTGGAGAGAGAACTGCAAGTGTACCAACCAAGGAGATTCGCCAAGGCCAT 2199
QY 1744 TGGCCAGGGGCCCTTGAAGCAACAGATGACAGGATGTGGGGAATTCAGTGGCACTGGA 1803
Db 2200 GATGAGGCTCAACAGGGCTTCATTTGCCACCAAGCTGTGCTGTGTCATCAGGAATGAG 2259
QY 1804 GGCCTTCCGACTACTCAGCG-----CCAAACCCATTAAGTGAACATCGG 1851
Db 2260 ATGTTGGCAGACAGCAACATGCTGTGGGGTTACCCGCCACTGTCCGGGTGACACCAAG 2319
QY 1852 TGCTATCCTTAGAGAACACAGTCCAGTGTGACTGCTGGACTGTGATCAAGTCCCTGCGAG 1911
Db 2320 TGCTTTATTTCTTCAAAACGACAAATCCACTGTGAGAGGAGCTGTACCAAGTCCAGCA 2379
QY 1912 GCTGGAAGACACAGCTGACATCGACACCGAGATTGAACCTGCAGACAGAAATTT 1971
Db 2380 GCGTGAAGGACCATAGGCTTACATTGATAAGAGATTGAAGTTCTCAAGATTAATTT 2439
QY 1972 AAGAACCTGAGGGAAGTCCGAGGTCACTCTGAAGAAAAAGCCGACCAAGAAATGTGACTGT 2031

Db 2440 AAGATTTAAGGAAGTGGGGACACCTAAAGAAAGAAACCTGAGAGTGTGCTGT 2499
QY 2032 CACAAATCAGCTTACACACCCACCAAGG-----CGGCTCAAGCACAGAGGCTCC 2085
Db 2500 GGTGACCAAGAGCTTATTAACAAGAGAGAGGTGTCAACAGACAGGAGAGCTAAAGAT 2559
QY 2086 AGTCTGATCTCTTTAGGAAGG-----CTGCAAGAGAGGACAAAGTGTGGCTGTG 2139
Db 2560 CACCTTTCACCTTCAAGGAGGCTGTGCTCCAGGAGGTGATAGCAAACTTCAGCTCTTC 2619
QY 2140 CGGAGCAGAGCGCAAGAGAACTCCGCAAGCTCTCAAGCGCTGCAGAACACAGAC 2199
Db 2620 AAGAGCATCGAG 2679
QY 2200 ACTGCAAGCATCGAGGCTTACGCTTTCACCCAGCAACCAAGCACTGGCAGAGCGG 2259
Db 2680 GAGTGTAGCTGCTGCTGCTTACCTGCTTTCAGCCATGACAAACCACTGGCAGAGTGC 2739
QY 2260 CTTTCTGACACTGGGCTTCTGCTGCTCACAGCGCCCAACAAATAACAGTACTGG 2319
Db 2740 CATTCTGGAATCTGGAATCTTCTGCTGCTGCAAGTTCTTAACAACTACTACTGG 2799
QY 2320 TGATGAGGACCATTAATGAGACTCAAAATTTCTCTTCTGTGAATTTGCAATGCTGCTTC 2379
Db 2800 TGTTTGGTACAGTCAACGAGAGCGCAAAATTTCTGTTTGTGAGTTTCTACTGCTTT 2859
QY 2380 CTAGTACTTTGATCTCAACAGACCCCTACCAGCTGATGAATGCACTGAGTGAACACTG 2439
Db 2860 CTGAATATTTGACATGAATACGGATCTTATCAGCTCACAATACAGTACACAGATA 2919
QY 2440 GACAGGAGTCTCTCAACAGGCTACACGCTACAGCTCATGAGCTGAGAGCTGCAAGGT 2499
Db 2920 GAACGAGCATCTTAATCAGCTACATACAGCTGATGAGCTCCGAGCTGCCAAGG 2979
QY 2500 TACAAGAGTGTAAACCCCGGACTCGAAACATGAGCTGGAGCTTAAAGATGAGAGAGC 2559
Db 2980 TATAAAGTGTCAACCAAGACCCAGAGCCTGGACATTTGGAGCTAAAGAGAGGAAAC 3039
QY 2560 TATGAGCAATACAG 2573
Db 3040 TATGACCCGACAG 3053

RESULT 8

AK028285 4428 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
DEFINITION enriched library, clone:323040B12 product:SULFATASE PP homolog
[Rattus norvegicus], full insert sequence.

ACCESSION

VERSION AK028285.1 GI:26390192

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

AUTHORS

Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

9279253

PUBMED

10349636

REFERENCE

2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

2049374

PUBMED

11042159

REFERENCE

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Tomita, M., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilmberg, L., Wyndham-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 50,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4428)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES
source

Location/Qualifiers
1. 4428
/organism="Mus musculus"

/mol_type="mrna"
/strain="CS7BL/6J"
/db_xref="FANTOM DB:3230401B12"
/db_xref="taxon:10090"
/clone="3230401B12"
/issue_type="head"
/clone_stage="14, 17 days embryo"
/dev_stage="14, 17 days embryo"
287. -2899
/note="unnamed protein product; SULFATASE FP homolog [Rattus norvegicus] (SPTR|AAL71906, evidence: PASTY, 96.4%ID, 100%length, match=2610)
putative"
/codon_start=1
/protein_id="BAC25858.1"
/db_xref="GI:26390193"
/translation="MKYSLMALLLAVLTGTLGLSLCSTVRSORPRGRIOERKINRPN
IILVLTDDDDVELGSLQVNMKTKIMEQGGATFTNAFTVTPMCCPSRSMGLTKGVHN
HNYVTNNCSPPSQMGMEPRTPAVYNTGYRTAFPGKILNYNGSYIIPGWREWL
GIUKNSRFTNYTCRNIKERHGFYADYFTDLITNESINFFKSKYKMYPRPIMMV
IISHAAPGDESAPOFSKLYPNASQHTPSYNAFPMKDHIMQYTGMPHPHMFPT
VLQKRLQTLMSVDDSVLRYNMLVSEGLDNTYIITADHGYYHIGQFLVKGMPY
DPDIRVFFIRGSPISVPOIVLNTDLAFTLIDLAGLSPSDVDGKSLKLDLLE
KPNRFTNKKAKINRDTFLVERGKFLRKESGKNIQSNHLPKVERKELCOQARY
QTACQPGNQWQCIEDTSGKLRHKCKGSPDLLTVHONARLYSGLDHDKCECHRD
SGTSSRSORKNQRFLENKGPYKPRVHTRQTRSLSEFEGEIVYDINLEELQV
LPPRSIAKHDEGHQGFIGHQAAGDIRNEMLADNNVAGLPATVTVTHKCLLENDT
ICEERLYOSAPAWDKHAYIDKEIVLQDKIKNLRVGRHLKCKKPEKQCGGDSYY
NKEGKRYOKLKLHSLHPFKEAAQVDSKQLQKFKERRRKKERKKRKGEGCSL
PGLTCTDHNHWQTAFWNLGSCFACSTSNNTVYCLRTVNETNFTLFCFEPATGLE
YFDMNTDYPQNLNTVHTVTRVSLINQLHQLMELRSCQGVKQCNPRPKSLDIGAEGGN
YDPRGQLWDGEG"

polyA_signal 4414. 4419
polyA_site 4428
/note="putative"
/note="putative"
BASE COUNT 1278 a 1049 c 1077 g 1024 t
ORIGIN

Query Match 31.5%; Score 1060.8; DB 11; Length 4428;
Best Local Similarity 65.7%; Pred. No. 5e-209;
Matches 1665; Conservative 0; Mismatches 812; Indels 57; Gaps 6;
QY 97 CTGCTGGGTGGAGCTCGCCCTTCCTGTGCGACCCACCGCTGAAAGGCGAGGTTTCAGAGG 156
DB 335 CTGCTGGGAGCGCTGTGTCCACCGTTCGTCCTCCAGAGGTTCCGAGAGAGGATACAGAG 394
QY 157 GACCGCAGGAACATCGCCCCAACATCATCTCTGGTGTGACCGAGCAGCAGGATGTGGAG 216
DB 395 GAACGAAAAAATCATCAGACCCCAATATATCTGTGTCTCATCTGACGACCCAGGATGTGGAG 454
QY 217 CTGGGTTCCATCAGGTGTGATGACAGCCCGCGCATCATGAGCAGCGCGGGGGGCGAC 276
DB 455 CTGGGTTCCCTCGAGTCATGACAGAGAGAGAGATCATGGAACAGGTTGGGGGCCACC 514
QY 277 TTTCATCAAGCCCTTGTGTACACACACCCATGTCTGCCCTCACGCTCTCTCCATCTCTCACT 336
DB 515 TTACCAATAGCTTTGTGACACAGCCCATGTCTGTCCATCAGCTCATCTCATCTCACT 574
QY 337 GCGAAGTAGTCTCAGCAACACCAACACCTTACCAACAATGAGAACTGCTCTCGCCCTCC 396
DB 575 GGGAGTAGTCTCATAAACCAATGTCTACCAACAATGAGAACTGCTCTCGCTCTCCCTCG 634
QY 397 TGGCAGGCAGCAGCAGAGAGCGCACCTTTGGCGTGTACCTCAATAGACACTGGCTACCGG 456
DB 635 TGGCAGGCAGCAGCAGAGCTCGGACCTTTGTGTGTATCTCAACAACACCGGCTACAGA 594
QY 457 ACAGCTTTCTTGGGAAGATCTTAAATGAATAACAAGGCTCTTACGTGCGCACCGGCTGG 516
DB 695 ACAGCTTTTGGAAAAATACCTCAATGAATAACATGCGAGCTATACCTCTCTCGATGG 754
QY 517 AAGGAGTGGTGGACTCTCTTAAACCTCCCGCTTTTAACTACACAGCTGTGTGCGGAAC 576

Db 755 CGAGAAATGCTCGGATTAATCAAGAATTCTCGTTTCTATATAATACACTGTTTGTGCAAC 814
QY 577 CGGGTGAAGAAAGACAGCGCTCCGACTACTCCAAGGATTACCTCACAGACTCATCACC 636
Db 815 GGCATCAGGAGAGCATGGAATTGATATGATGAAGGATTACTTCAGACTTAATCACT 874
QY 637 AATGACAGGTGAGTCTTTCGCGAGCTCCAAAGAGATGTACCGGCACAGCGCAGTCTC 696
Db 875 AACGAGAGCATTAATTACTTCAAAATGTCTAAGAAATGTATCCCATAGGCCCATATG 934
QY 697 ATGGTCATCAGCCATGACAGCCCCCAGCGCCCTGAGGATTTCAGCCCAATATTCAGC 756
Db 935 ATGGTATCAGCCATGACAGCACATGCGCCCTGAAGACTCGGCCCAAGTTCATCAAA 994
QY 757 CTCCTCCCAACGCACTCTCAGCACATCAGCGGAGCTACAATAGCGCCCAACCGGAC 816
Db 995 CTGTACCCCAATGCTCCCAACACATAACACCCAGCTATACTATGACCAAGCATGGAC 1054
QY 817 AAACATGATCATGGCTACAGCGGCGCCCTGAGGCCATCCATGCAATTCACCAAC 876
Db 1055 AAACATGATCATGGCTACAGCGGCGCCCTGAGGCCATCCATGCAATTCACCAAC 1114
QY 877 ATGCTCCAGCGAAGCGCTTCAGACCCCTCATGTGCGTGGAGCACTCCATGGAGCGATT 936
Db 1115 GTCTCCAGCGAAGCGCTTCAGACCTCTGATGTGAGTTCAGTTCAGACTCTGTGAGAGTTG 1174
QY 937 TACAATGCTGTTGAGAGCGGCGAGCTGGACACAGTATACATGATACACCGCGAC 996
Db 1175 TACAATGCTGTTGAGAGCGGCGAGCTGGACACAGTATACATGATACACCGCGAC 1234
QY 997 CACGGTTACCATCGGCCAGTTTGGCTGTGGAAGGGAATCCATGCAATGAGTTT 1056
Db 1235 CACGGTTACCATCGGCCAGTTTGGCTGTGGAAGGGAATCCATGCAATGAGTTT 1294
QY 1057 GACATCAGGGTCCGGTTCTAOTGAGGGGCCCCAAAGTGGAGCGGCTGTCTGAATCCC 1116
Db 1295 GACATCAGGGTCCGGTTCTAOTGAGGGGCCCCAAAGTGGAGCGGCTGTCTGAATCCC 1354
QY 1117 CACATCGTCTCAACATTCAGCTGCGCCGCCCAATCTTGGACATTCAGAGGCTGGACATA 1176
Db 1355 CAGATTGTTCTGAACATCGACCTAGCTCTACCATCTTGATATGATGTTGTTGACTCT 1414
QY 1177 CTGCGGATATGAGCGGGAATTCATCTCAAGTGTGGAACGAGCGCGCGGTGAAT 1236
Db 1415 CTTCTGATGTGATGTTAAATCTGTCTCAAACTCTAGACCTGGAAGCCGGTAAAC 1474
QY 1237 CGGTTTCACTTGAAGAGAGATGAGGTCTGGCGGACTCTTCTTCTGTTGAGAGAGC 1296
Db 1475 AGTTTTCGACAAACAGAGAGCCAAATCTGGCGTGATACATTCCTAGTGAAGAGGG 1534
QY 1297 AAGCTGCTACAAAGAGAGCAATGACAGGTGGAGCGGCCAGGAGAGAACTTTCTGCC 1356
Db 1535 AAATTTTACGGAAGAGAGGAATCCGCAAGATATCCAAAGTCAACCACTTGCCC 1594
QY 1357 AAGTACAGCGTGTGAGAGACCTGTGTAGCTGTGAGTACAGACCGCGTGTGAGAG 1416
Db 1595 AAATATGAGAGGTCAAGAACTGTGCCAGAGCCAGATACAGACAGCGCTGTGAGAG 1654
QY 1417 CTGGGACAGAGTGGAGTGTGTGAGGAGCGCCAGCGGAAGCTGAAGCTGCATAAGTGC 1476
Db 1655 CTGGGACAGATGGAGTGTGTGAGGAGCGCCAGCGGAAGCTGAAGCTGCATAAGTGC 1714
QY 1477 AAGGGCC-----CAGTGGCTGGGCGCAGCAGAGCGCTCTCCAA 1517
Db 1715 AAGGACCCAGCGACCTGTCTACGGTCCATCAGATGACAGCAACCTCTACTCTCGCGGA 1774
QY 1518 CCTCGTGGCCCAAG--TACTACGGGAGGCGAGCGAGGCTGCACTGTGAGAGCGGAGC 1575
Db 1775 TTGCATGACAAAGACAAAGAGTGCCATTGTAGGACTCTGGCTATCGCTCGAGCAGAGC 1834
QY 1576 TACAAGCTCAGCTGCGCGGACCGCGGAAAAAATCTTTCAAGAGAGTACAAGGCCAGC 1635
Db 1835 CAGAGAAAGAAATCAAGGCGAGTCTCTGAGGAAACAGGGAACACCAAGAGTATAAGGCCAGA 1894

RESULT 9
AK008108
LOCUS

QY 1636 TATGTCCGAGTGTGCTCCATCCGCTCAGTGGCCATCGAGTGGAGCGGAGGTGTACCAC 1695
Db 1895 TTGTGCAACACCCGCGAGACTCGTCCCTGTCTGTGAATTCGAAGTGAATATATGAC 1954
QY 1696 GTAGGCTGGGTGATGCGCCCGAG-----CCCCGAACTCACCAGCGGCAC 1743
Db 1955 ATCAACTTGGAAAGAAAGAACTGCAAGTGTTCACCAAGGAGCATTCGCAAGGCCAT 2014
QY 1744 TGCCAGGCGGCGCTGAGGACCAAGATGACAGGATGTGGGGAATTCAGTGGCACTGGA 1803
Db 2015 GATGAGGTTCACCAAGGCTTCATTGGCCACCAGGCTGCTGCTGGTGACATCAGGAATGAG 2074
QY 1804 GGCCTTCCGCACTACTAGCCG-----CCAAACCCCATTAAGTGTACATCGG 1851
Db 2075 ATGTTGGCAGACAGCAACAATGCTGTGGGTTTACCGCCACTGTCCGGGTGACACACAAG 2134
QY 1852 TGTCTACATCTAGAGAAACACACAGTCCAGTGTGACCTGGACCTGTAAGTTCCTTCGAG 1911
Db 2135 TGTCTTATTTCTTCCAAACGACACATCTGCTGAGAGGAGCTGTACCAGTCAGCCAGA 2194
QY 1912 GCTGTGAAGAGACCAAGCTGCACATCGACCAAGATTTGAAACCTTCGAGAACAAATTT 1971
Db 2195 GCTGTGAAGAGACCAATAGGCTTACATTTGATAAGAGATTGAAAGTTCTACAGATAAATTT 2254
QY 1972 AGAACCTCAGGAGAGTCCGAGGTCACTGAGAAAGAAAGCGCCAGAGAAATGTGACTGT 2031
Db 2255 AGAATTTTAAAGGAGTGAAGGACACCTTAAGAAAGAAAGAACTTGAGAGTGTGGCTGT 2314
QY 2032 CACAAATCAGCTTACCACACCCAGCACAAGG-----CCGCTTCAAGCAGAGAGCTCC 2085
Db 2315 GGTGACCAAGCTATTACAAACAAAGAGAAAGGTGTCAAACGACAGAGAGAGCTAAAGAT 2374
QY 2086 AGTCTCATCTTTCAGGAAGG-----CCTGCAAGAGAGAGCAAGTGTGGCTGTG 2139
Db 2375 CACCTTTCACCTTCAAGGAGGCTGTGCTCCAGAGGTTGATAGCAAACTTCAGTCTTTC 2434
QY 2140 CGGAGCAGAGCGCAGAAAGAACTCCGCAAGCTGTCAAGCGCTTCAGAACCAACGAC 2199
Db 2435 AAGGACATCGGAGGAGGAAAGAGAGAGAGAGAGAAACCGCAGAGAGAGAGAG 2494
QY 2200 AGTGCAGATGCGAGGCTCTACGTGCTTACCGACGACCAACCGACTGCGCAGAGCGG 2259
Db 2495 GAGTGTAGCTGCTGCGCTTACCTGCTTACCCATGACAAACCACTGGCAGAGTGGC 2554
QY 2260 CTTTCTGACACTGGGGCTTTCTGTGCTGCAACAGCGCCAAACAAATACACGACTGCTG 2319
Db 2555 CCATTTCTGAACTTGGGATCTTCTGTGCTGCACAAGTTCTTAAACAATACCTACTG 2614
QY 2320 TGATGAGGACCATCAATGAGACTCAAAATTTCTTCTGTGAAATTTGCAACTGCTTC 2379
Db 2615 TGTGCTGACAGTCAACGAGAGCGCAAAATTTCTGAGTTTGTGAGTTTGTCTACTGCTTT 2674
QY 2380 CTAGAGTACTTTGATCTCAACACAGACCCCTACAGCTGATGAATGCAATGCAACACTG 2439
Db 2675 CTGGAATATTTGACATGAATGAGTCTTATCAGCTCACAATACAGTACACAGTA 2734
QY 2440 GACAGGAGTGTCTCAACAGCTTACAGTACAGTCACTGAGTCACTGAGGAGCTGCAAGGT 2499
Db 2735 GAACGGAGCATCTTAATCAGCTTACATACAGTGTGAGTCTGAGGCTGCAAGGG 2794
QY 2500 TACAAGCTGTAAACCCCGGACTCGAAACATGAGCTTGGAGCTTAAAGATGAGGAGC 2559
Db 2795 TATAACAGTGCACCAAGACCCCAAGAGCTTGGAGCTTAAAGAGAGAGAAAC 2854
QY 2560 TATGAGCAATACAG 2573
Db 2855 TATGACCCGACAG 2868

AK008108 1803 bp mRNA linear HTC 05-DEC-2002

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tsgama, M., Tsgawa, A., Takanashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physiol and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. CDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTGAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOUR.

Location/Qualifiers
1. 1803
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:2010004N24"
/db_xref="MGI:1904769"
/db_xref="taxon:10090"
/clone="2010004N24"
/sex="male"
/tissue type="small intestine"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stages="adult"
715. 1164
/note="unnamed protein product: putative weakly similar to N-ACETYLGLUCOSAMINE-6-SULFATASE [Coturnix coturnix] (SPRIF090XB6, evidence: FASTA, 64.2%ID, 99.5%length, match=2601)"
/codon_start=1
/protein_id="BAB25464.1"
/db_xref="GI:12842086"
/db_xref="MGI:1919293"
/translation="MPSGLCTHDHWHQTPALWTLGPFCACTSANNTYWCRLPTINE THNPLCEFAFTGFLYEDLSITDPYOLMNAVNTLDRDLNLQLHVLQMLRSGKGYQCQN PRTFMNLDLGRDGGSEYEQYQFORKWPKEPKRPSKSLQGLWEGWEG"
1778. 1783
/note="putative"
1803
/note="putative"
481 a 512 c 467 g 343 t

h 25.5% Score 856.4; DB 11; Length 1803;
Similarity 87.0%; Pred. No. 1.1e-166;
41; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

1 CCAGAGTCGTCCTCCGCTCAGTGGCCATCGAGTGCACGCGAGGTGTACCACTAGG 1700
7 CCTCGGCTCTCCATCCGTTCCGTGGCCATCGAGTGCACGCTGAGATATACCACTAGG 206
CCTGGGTGATCGCCCCAGACCCCAAAACCTCACCAGGCGGACCTGGCCAGGGGCCCTGA 1760

207	CTTGGATAC	TGTGCTCAGCCCGCAACCTTAGCAAGCCGACACTGGCCAGGGGCGCCTCGA	266
1761	GGACCAAGATGACA	AGAGTGGTGGGACATTACGTGGCACTGGAGGCGCTTCCCGACTACTC	1820
267	AGACCAAGATGACA	AGAGTGGTGGGACATTACGTGGTACTAGTGGCTTCCAGATTATTC	326
1821	AGCGCCCAACCCCAT	TAAAGTGCACATCGGTCTACATCTCTAGAGACACACACAGTCCA	1880
327	TGCCCCCAATCCCAT	CAAGTGAACCATCGGTGCTACATCTTGGAGATGACACAGTCCA	386
1881	GTGTGACCTTGGAC	CTGTACAAAGTCCCTGCAGGCGCTGGAAAGACACAAAGGTGCACATCGA	1940
387	GTGCGACTTGGAC	CTGTACAAAGTCCCTGCAGGCTTTGGAAGACACACAAGCTGCACATCGA	446
1941	CCACGAGATTGAAC	CCCTGCAGAACCAAAATTAAGAACCTGAGGAGAGTCCGAGGTCACTT	2000
447	CCATGTGATGCAAA	ACCCCTGCAGAACCAAAATTAAGAACCTTGCAGAAAGTTCAGGGGTCACTT	506
2001	GAAGAAAAGCGGC	CAAGAAATGTGACTGTGTCAAAAATCAGCTTACACACCCACAGCAAA	2060
507	GAAAGAGAGCGAG	CCGGAAGATGTGACTGTGCCATAGAAATCAGTTACACAGGCCACACAA	566
2061	AGCGCGGCTCAAG	CAAGAGGCTCCAGTCTGCATCTCTTTCAGGAAGGGCTGCGAAGAA	2120
567	AGCGCGTCTCAAG	CACAAAGGCTCCAGGCTGCACCTTTTCAGGAAGGGTCTGCGAGGAA	626
2121	GGACAAAGTGTGG	CTGTTGGCGAGCAGAAAGCGCAAGAGAAACTTCGCAAGCTGTCTCAA	2180
627	GGACAAAGTGTGG	CTGTTGGCGAGCAGAAAGCGCAAGAGAAACTTCGCAAGCTGTCTCAA	686
2181	CGCGCTGAGAAAC	AAAGACACGTGCGAGCATGCCAGGCTTCAGTGTTCACCCACGACAA	2240
687	ACGGCTGCAGAA	CAACGATACGTGCGAGCATGCCAGGCTTCAGTGTTCACCCACGACAA	746
2241	CCAGCACTGGCAG	ACGCGCTTCTTCGGACACTGGGGCTTCTGTGCTGCGACACAGCGC	2300
747	CCACCACTGGCAG	ACGCGCCACTCTGAGAGCTTGGGGCGCTTTCGCGCTGCGACAGCGC	806
2301	CAACAATAACACG	TACTGTGTGCATGAGGACCAATCAATGAGACTCAAAATTTCTCTTCTG	2360
807	CAACAACAACACG	TACTGTGTGCTTGAGGACCAATAAATGAGACCCACAACTTCCTCTCTG	866
2361	TGAATTTGCACTG	GCTTCCTGAGTACTTTGATCTCAACACAGACCCCTACAGCTGAT	2420
867	CGAATTTGCAACC	GCGTTCATGAATACTTTGACCTCAGTACAGACCCCTACAGCTGAT	926
2421	GAATGCAGTGAACA	CACTGGCAGGGATGTCTCTCAACAGCTTACACGTACAGCTCATGGA	2480
927	GAAACGCGTGAA	CACACTGGACAGGACGTCTTAAACCACTGCACGTGCAGCTCATGGA	986
2481	GCTGAGGAGCTGCA	AGGGTTACAAGCAAGTGTAAACCCCGACTCGAAACATGGACCTGGG	2540
987	GCTAAGGAGCTGT	AAAGGCTACAAAGCAGTGCAACCCCGGACCCGCAACATGGACCTGGG	1046
2541	ACTTAAAGATGGAG	GAAGCTATCAGCAATACAGGCAAGTTTCAGCGTCCAAAAGTGGCCAGA	2600
1047	GCTTAGAGACGAG	GAAGCTATGAAACAATACAGGCAAGTTTCAGCGTCCAAAATGGCCAGA	1106
2601	AATGAAGAGACTT	CTTCCAAATCACTGGGACAACTGTGGGAAGGCTGGGAAGGTTAAGA	2660
1107	AATGAAGAGACTT	CTTCTCCAAATCACTGGGACAGCTATGGGAAGGTTGGGAAGGCTAAGC	1166
2661	AACAAACAGGTTG	ACCTCCAAAACATAGAGGCATCAGCTGACAGGCAATGAAA	2720
1167	GGCCATAGAGAG	AGAACCTTCCAAACACAGGGGCGCTCGTGTGCTGCCAGGCCATGCAA	1226
2721	AA	2722	
1227	AA	1228	

AL537332 1201 bp mRNA linear EST 31-MAY-2003
 LOCUS AL537332 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 DEFINITION CSODF025YE15 5-PRIME, mRNA sequence.
 ACCESSION AL537332
 VERSION
 KEYWORDS
 SOURCE EST
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLES Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12800825.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9923.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODF025AC08Qp1&cluster=9923.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODF025AC08Qp1.
 FEATURES
 source Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODF025YE15"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: PCWSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the PCWSPORT 6
 vector. Library was not normalized."
 BASE COUNT 296 a 319 c 335 g 193 t 58 others
 ORIGIN
 Query Match 24.9%; Score 838.2; DB 9; Length 1201;
 Best Local Similarity 95.1%; Pred. No. 6.5e-163;
 Matches 904; Conservative 4; Mismatches 37; Indels 6; Gaps 4;
 QY 1282 TTGGTGGAGAGGGCAAGTGCTTACAAAGAGAGACAATGACAAGGTGGACGCCAGGAG 1341
 Db 20 TTGTGACAAAAGCAGGTGGTACCGGTCCGGATTCGGGATGTGGACGCCAGGAG 79
 QY 1342 GAGAACTTTCTGCCCAAGTACACAGCGTGTGAAGGACCTGTGTACAGCGTGTGTAGTACCAG 1401
 Db 80 GAGAACTTTCTGCCCAAGTACACAGCGTGTGAAGGACCTGTGTACAGCGTGTGTAGTACCAG 139
 QY 1402 ACGCGGTGTGAGCAGCTGGGACAGAGTGTGCAGTGTGTGAGGACGCCACGGGGAGCTG 1461
 Db 140 ACGGCGTGTGAGCAGCTGGGACAGAGTGTGCAGTGTGTGAGGACGCCACGGGGAGCTG 199
 QY 1462 AAGCTGCATAAGTGCAGGGGCCCCCATCGCGTGGCGGAGCAGAGCCCTTCTCAAACCTC 1521
 Db 200 AAGCTGCATAAGTGCAGGGGCCCCCATCGCGTGGCGGAGCAGAGCCCTTCTCAAACCTC 259
 QY 1522 GTGCCCAAGTACTACGGGAGGGGACGAGGGCTGCACCTGTGCACAGCGGGGACTACAAG 1581
 Db 260 GTGCCCAAGTACTACGGGAGGGGACGAGGGCTGCACCTGTGCACAGCGGGGACTACAAG 319
 QY 1582 CTCAGCTGCGCGGAGCGCGGAAAAAAGCTTTCAGAAAGAGTACAAGGCCAGCTATGTC 1641
 Db 320 CTCAGCTGCGCGGAGCGCGGAAAAAAGCTTTCAGAAAGAGTACAAGGCCAGCTATGTC 379

```

QY 1642 CGCAGTCGCTCCATCCGCTCAGTGGCCATCGAGTGGAGCGGAGGGTGTACACCTAGGC 1701
Db 380 CGCAGTCGCTCCATCCGCTCAGTGGCCATCGAGTGGAGCGGAGGGTGTACACCTAGGC 439
QY 1702 CTGGGTGATGCGCGCCAGCCGAGGACCTTCCAAAGCGGCACTGGCCAGGGGCCCTCGAG 1761
Db 440 CTGGGTGATGCGCGCCAGCCGAGGACCTTCCAAAGCGGCACTGGCCAGGGGCCCTCGAG 499
QY 1762 GACCAAGATGACAGAGATGGTGGGACCTTCACTGGGCACTGGAGGGCTTCCCGACTACTCA 1821
Db 500 GACCAAGATGACAGAGATGGTGGGACCTTCACTGGGCACTGGAGGGCTTCCCGACTACTCA 559
QY 1822 GCGGCCAACCCCATTAAGTGTACACATCGTGTCTACATCTCTAGAGAACGACACATGCCAG 1881
Db 560 GCGGCCAACCCCATTAAGTGTACACATCGTGTCTACATCTCTAGAGAACGACACATGCCAG 619
QY 1882 TGTGACCTGGACCTGTACAAAGTCCCTGAGGCTGGAAAGACCAACAGCTGCACATCGAC 1941
Db 620 TGTGACCTGGACCTGTACAAAGTCCCTGAGGCTGGAAAGACCAACAGCTGCACATCGAC 679
QY 1942 CACGAGATTGAACCTCGACAGACAAATTAAGAACCTTGAGGAGTCCGAGGTCACCTG 2001
Db 680 CACGAGATTGAACCTCGACAGACAAATTAAGAACCTTGAGGAGTCCGAGGTCACCTG 739
QY 2002 AAGAAAGAGCGCCAGAGAAATGTGACTGTCTACAAATATCAGTACACACCCAGCACAAA 2061
Db 740 AAGAAAGAGCGCCAGAGAAATGTGACTGTCTACAAATATCAGTACACACCCAGCACAAA 799
QY 2062 GCGCGCTCAAGCAGACAGGCTCGAGTGTGATCTTTCAGAGAGGGCTCGAAGAGAG 2121
Db 800 GCGCGCTCAAGCAGACAGGCTCGAGTGTGATCTTTCAGAGAGGGCTCGAAGAGAGAG 859
QY 2122 GACAAAGTGTGCTGTGTGGGAGCAGAGAGCGCAAGAGAAATCCGCAAGCTGCTCAAG 2181
Db 860 GACAAAGTGTGCTGTGTGGGAGCAGAGAGCGCAAGAGAAATCCGCAAGCTGCTCAAG 916
QY 2182 CCGCTGAGAACACGACAGCTGCGAGAGTGGAGGCTTCACTGCTTCAACC 2232
Db 917 CCGCTGAGAACACGACAGCTGCGAGAGTGGAGGCTTCACTGCTTCAACC 964

RESULT 11
AL574439/c
LOCUS 1201 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION clone CS0DI058YD10 3-PRIME, mRNA sequence.
AL574439
AL574439.2 GI:31312757
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12934646.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9923.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI058YD105NP1&cluster=9923.r. Contact :
Feng Liang Email : fliang@lifetech.com URL : Corporation 1600
http://fulllength.invitrogen.com/ Invitrogen
Paradise Avenue Genoscope sequence ID : CS0DI058YD105NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
FEATURES
source

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI058YD10"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 273 a 253 g 334 t 100 others
ORIGIN
Query Match 24.7%; Score 829.4; DB 9; Length 1201;
Best Local Similarity 92.5%; Pred. No. 4.3e-161;
Matches 864; Conservative 40; Mismatches 21; Indels 9; Gaps 4;
QY 2391 TGATCTCAACACAGAGACCCCTTACCAGCTGATGAATGAGTGAACACACACTGGACAGGAGTGT 2450
Db 934 TTATYTCACACAGAGACCCCTTACCAGCTGATGAATGAGTGAACACACACTGGACAGGAGTGT 875
QY 2451 CCTCAACCAAGTACAGTACAGCTCATGAGTGAAGAGCTGCAAGGGTTACAAGCAGTG 2510
Db 874 CCTCAACCAAGTACAGTACAGCTCATGAGTGAAGAGCTGCAAGGGTTACAAGCAGTG 815
QY 2511 TAACCCCGGAGCTCGAAACATGAGACCTGGAGCTTAAGATGAGGAGGCTATGAGCAAA 2570
Db 814 TAACCCCGGAGCTCGAAACATGAGACCTGGAGCTTAAGATGAGGAGGCTATGAGCAAA 755
QY 2571 CAGGCAAGTTTACGCTGCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGG 2630
Db 754 CAGGCAAGTTTACGCTGCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGG 595
QY 2631 ACACCTGTGGAGAGGCTGGAGAGTTAAGAAACACAGAGGTGGAGCTTCCAAACATAG 2690
Db 694 WCAAYTGTGGAGAGGCTGGAGAGTTAAGAAACACAGAGGTGGAGCTTCCAAACATAG 635
QY 2691 AGGCATCACTGACTGCACAGGCAATGAAACCACTGTGGG-TGATTTTCCAGC-AGACCT 2748
Db 634 AGGCATCACTGACTGCACAGGCAATGAAACCACTGTGGGTTGATTTCCAGCTAGACT 575
QY 2749 GTGCTATTGCGCAGAGGCTCGAGAAAGCAAGACGCACTCTCAGTCAACATGACAGATT 2808
Db 574 KTGTCTTTGGAGTATACCTCACCAGCTGCAAAATGCAATTTTTCGTATCAAAAGCCACC 515
QY 2809 CTGAGAGTAACACAGCAGAGCAGAGATACTTCAAGAGTCCATTTTGGCCCTGCTTT 2868
Db 514 CTGAGAGTAACACAGCAGAGCAGAGATACTTCAAGAGTCCATTTTCCCTCTCTTT 455
QY 2869 TGCTTTGGATTATACCTCACCAGCTGCAAAATGCAATTTTTCGTATCAAAAGTCAACC 2928
Db 454 TGCTTTGGATTATACCTCACCAGCTGCAAAATGCAATTTTTCGTATCAAAAGCCACC 395
QY 2929 ACTTACCTTCCCGCAGAGCTCACAAGAGAAACGAGAGAGCGAGAGAGAGATTTC 2988
Db 394 ACTTACCTTCCCGCAGAGCTCACAAGAGAAACGAGAGAGCGAGAGAGAGATTTC 335
QY 2989 TTGGAATTTCTCCCAAGGGCGAAAGTCAATGGAATTTTAAATCATAGGGGAAAGCAG 3048
Db 334 TTGGAATTTCTCCCAAGGGCGAAAGTCAATGGAATTTTAAATCATAGGGGAAAGCAG 275
QY 3049 TCTGTTCTAAATCTCTTATTTCTTTGTTGTTGTCACAAA-----GAAGGAACTAAGAA 3102
Db 274 TCTGTTCTAAATCTCTTATTTCTTTGTTGTTGTTGTCACAAA-----GAAGGAACTAAGAA 215
QY 3103 GCAGGAC-AGAGGCAACGTGGAGAGGCTGAAACACAGTGCAGAGAGCTTTGCAATGAGTC 3161
Db 214 GCAGGAC-AGAGGCAACGTGGAGAGGCTGAAACACAGTGCAGAGAGCTTTTACATGAGTC 155
QY 3162 AGTAGCAGAAAGAGATGACATTTACCTAGCATTATAAACCTGTTGCTCTTGAAGAA 3221
Db 154 AGTAGCAGAAAGAGATGACATTTTCCCTAGCATTATAAACCTGTTGCTCTTGAAGAA 95
QY 3222 CTGCTTCTATATATATATGAGTATTTTACATGTAATCAACATGGGAACTTTTAGGGA 3281

```

```
Db 94 CTGCTTCATTCTATATATGTTACTATTATACATKTAATCAACATGGGAGCTTTTGGGCGN 35
QY 3282 ACCTAATAGAAATCCCAATTTTCAGGAGTGCTG 3315
Db 34 NCCTAATAGTATCCCAATTTTCAGGAGTGCTG 1

RESULT 12
BU956086 948 bp mRNA linear EST 21-OCT-2002
LOCUS AGENCOURT_10621456 NIH_MGC_107 Homo sapiens cDNA clone
DEFINITION IMAGE:6729576 5', mRNA sequence.
ACCESSION BU956086
VERSION BU956086
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3052 row p column: 23
High quality sequence stop: 552.
Location/Qualifiers
1. .948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6729576"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."
BASE COUNT 294 a 201 c 226 g 222 t 5 others
ORIGIN
Query Match 24.2%; Score 812.8; DB 13; Length 948;
Best Local Similarity 97.9%; P-red. No. 1.2e-157;
Matches 853; Conservative 0; Mismatches 15; Indels 3; Gaps 3;
QY 2484 GAGGAGCTGCAAGGTTACAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGGACT 2543
Db 1 GAGGAGCTGCAAGGTTACAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGGACT 60
QY 2544 TAAAGATGAGGAGGAGCTATGAGCAATACAGGAGCTTTTCAGCGTCCGAAAGTCCGAGAAAT 2603
Db 61 TAAAGATGAGGAGGAGCTATGAGCAATACAGGAGCTTTTCAGCGTCCGAAAGTCCGAGAAAT 120
QY 2604 GAAGAGACCTTTTCCAAATCACTGGGACAACTGTGGGAGGCTGGGAAGGTTAAGAAAC 2663
Db 121 GAAGAGACCTTTTCCAAATCACTGGGACAACTGTGGGAGGCTGGGAAGGTTAAGAAAC 180
QY 2664 AACAGAGGTTGAGCTCCAAAACATAGAGGCACTCACTGCTGACAGGCAATGAAAC 2723
```



```

FEATURES
source
High quality sequence stop: 634.
Location/Qualifiers
1..1066
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5805439"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: Breast; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT      325 a      286 c      258 g      197 t
ORIGIN
Query Match      23.8%; Score 799; DB 13; Length 1066;
Best Local Similarity 90.2%; Pred. No. 8.4e-155;
Matches 968; Conservative 0; Mismatches 40; Indels 65; Gaps 8;

QY 1983 GGAAGTCGCGAGTCACTGAGCAAGAAAGCGGCAGAGATGTGACTGTCAAAATCAG 2042
DB 1 GGAAGTCGCGAGTCACTGAGCAAGAAAGCGGCAGAGATGTGACTGTCAAAATCAG 60
QY 2043 CTACACACCCAGCAAAAGCGCGCTCAAGCACAGAGGCTCCAGTCTGCATCTTTTCAG 2102
DB 61 CTACACACCCAGCAAAAGCGCGCTCAAGCACAGAGGCTCCAGTCTGCATCTTTTCAG 120
QY 2103 GAAGGCGCTGCAAGAGAGGCAAGGTGTGCTGTGGGAGCAGAGCCAGAGAA 2162
DB 121 GAAGGCGCTGCAAGAGAGGCAAGGTGTGCTGTGGGAGCAGAGCCAGAGAA 180
QY 2163 ACTCGCGAGTGTCTCAAGCGCTCCAGAAACACACACGTCGACGATGCCAGGCTCAC 2222
DB 181 ACTCGCGAGTGTCTCAAGCGCTCCAGAAACACACACGTCGACGATGCCAGGCTCAC 240
QY 2223 GTGCTTACCCAGACACACGACATCGGAGACGCGGCTTCTTGGACACTGGGGGCTTT 2282
DB 241 GTGCTTACCCAGACACACGACATCGGAGACGCGGCTTCTTGGACACTGGGGGCTTT 300
QY 2283 CTGTGCTGCACAGCGCCACAAATAACACGCTACTGTGTGATGAGGACCATATGAGAC 2342
DB 301 CTGTGCTGCACAGCGCCACAAATAACACGCTACTGTGTGATGAGGACCATATGAGAC 360
QY 2343 TCACAAATTCCTTCTTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAACAC 2402
DB 361 TCACAAATTCCTTCTTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAACAC 420
QY 2403 AGACCCCTACGAGTGAATGAGTGAACACACACGAGGATGCTCTCAACGAGCT 2462
DB 421 AGACCCCTACGAGTGAATGAGTGAACACACACGAGGATGCTCTCAACGAGCT 480
QY 2463 ACACGTACAGCTCATGAGAGTGAAGGTTTACAAGCAGTGTAAACCCCGGAC 2522
DB 481 ACACGTACAGCTCATGAGAGTGAAGGTTTACAAGCAGTGTAAACCCCGGAC 540
QY 2523 TCGAAACATGACCTGGGACTTAAGATGAGGAGGATGATGAGCAATACAGGAGTTTCA 2582
DB 541 TCGAAACATGACCTGGGACTTAAGATGAGGAGGATGATGAGCAATACAGGAGTTTCA 591
QY 2583 GGTGCGAAAGTGGCGCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACACTGTGGGA 2642
DB 592 -----GGGACAACTGTGGGA 606
QY 2643 AGGCTGGGAAGGTTAAGAAACAAACAGAGGTGAGCTTCCAAATAATAGAGCATCACCTG 2702
DB 607 AGGCTGGGAAGGTTAAGAAACAAACAGAGGTGAGCTTCCAAATAATAGAGCATCACCTG 666

```

2703 ACTGCACAGCAATGAAACACCACTGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAG 2762
 667 ACTGCACAGCAATGAAACACCACTGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAG 726
 2763 GAGGCTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATTAACCA 2822
 727 GAGGCTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATTAACCC 786
 2823 GCAGGAGCAGAG-ATAACTTCAGGAAGT-CCATTTTGGCCCTGCTTTTG-CTTTCGATT 2879
 787 GCAGGAGCAGAGATTAATCTCAGGAATCCATTTTGGCCCTGCTTTTGCCCTTGGATT 846
 2880 ATACCTCAGCAGCTGCACAAAATG--CATTTTCTGATCAAAAAGTCAACCTAACCT 2937
 847 ATACCTCAGCAGCTGCACAAAATGCAATTTTTCGAATCAAAAAGTCAACCTAACCT 906
 2938 CCCCCAGAA--GCTCAAAAGGAACGAGAGAGAGCGAGAGAGAGA--TTTCCTTGA 2993
 907 CCCCCAAAAGCTTCCAAAGGAACCGAAAACCAACCAAAAATAATTTCTCTTGA 966
 2994 AATTTCTCCCAAGGCGCAAGTC--ATTGGAATTTTAAATCATAGGGGAAA 3044
 967 AATTTCTCCCAAGGCGCAAGTCATTGGGAATTTTAAATCTAGGGGAAA 1019

BU956560 885 bp mRNA linear EST 21-OCT-2002
 AGENCOURT 10615432 NIH_MGC 107 Homo sapiens cDNA clone
 IMAGE:6730171 5', mRNA-sequence.
 BU956560
 BU956560.1 GI:24186132
 EST.
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 885)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 Unpublished
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
 Plate: LLCM3054 row: i column: 18
 High quality sequence stop: 732.
 Location/Qualifiers
 1..885
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6730171"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_107"
 /note="Organ: breast; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 291 a 183 c 218 g 192 t 1 others
 ORIGIN
 Query Match 23.7%; Score 798.6; DB 13; Length 885;

Best Local Similarity 98.1%; Pred. No. 1e-154;
Matches 829; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 2484 GAGGAGCTGCAAGGGTTAACAAGAGGTAAACCCCGGACTCGAAACATGAGACCTGGGACT 2543
DB 1 GAGGAGCTGCAAGGGTTAACAAGAGGTAAACCCCGGACTCGAAACATGAGACCTGGGACT 60

QY 2544 TAAAGATGAGGAGGATGATGAGCAATACAGGAGGATTTTCAGGCTGCAAAAGTGCCAGAAAT 2603
DB 61 TAAAGATGAGGAGGATGATGAGCAATACAGGAGGATTTTCAGGCTGCAAAAGTGCCAGAAAT 120

QY 2604 GAAGAGACCTTTTCCAAATCACTGGGCAACTGTGGGAGGCTGGGAGGTTTGAAGAAC 2663
DB 121 GAAGAGACCTTTTCCAAATCACTGGGCAACTGTGGGAGGCTGGGAGGTTTGAAGAAC 180

QY 2664 ACAGAGGTTGGACCTCCAAACATAGAGGATCACTGCTGCTGACAGGATGAAAC 2723
DB 181 ACAGAGGTTGGACCTCCAAACATAGAGGATCACTGCTGCTGACAGGATGAAAC 240

QY 2724 CATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGAGGCTTGAGAAAGCAAGCAC 2783
DB 241 CATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGAGGCTTGAGAAAGCAAGCAC 300

QY 2784 GCACCTCTCAGTCAACATGACAGATTTCTGAGGATTAACAGCAGGAGGACAGATTAATTCA 2843
DB 301 GCACCTCTCAGTCAACATGACAGATTTCTGAGGATTAACAGCAGGAGGACAGATTAATTCA 360

QY 2844 GGAAGTCCATTTTGGCCCTGCTTTTGGCTTTTGGATTAATACCTCACAGCTGCACAAAATG 2903
DB 361 GGAAGTCCATTTTGGCCCTGCTTTTGGCTTTTGGATTAATACCTCACAGCTGCACAAAATG 420

QY 2904 CATTTTTCGTATCAAAAAGTACCACTAACCCCTCCAGAGGCTCAAGAGGAAACG 2963
DB 421 CATTTTTCGTATCAAAAAGTACCACTAACCCCTCCAGAGGCTCAAGAGGAAACG 480

QY 2964 GAGAGAGCAGGAGAGAGATTTCTTGGAAATTTCTCCAGAGGCGAAAGTCATTGAA 3023
DB 481 GAGAGAGCAGGAGAGAGATTTCTTGGAAATTTCTCCAGAGGCGAAAGTCATTGAA 540

QY 3024 TTTTAAATCATAGGGGAAAGAGCTGCTGTTCTTAATCTCTTATTTCTTTTGGTTTCTC 3083
DB 541 TTTTAAATCATAGGGGAAAGAGCTGCTGTTCTTAATCTCTTATTTCTTTTGGTTTCTC 600

QY 3084 ACAAAGAGAACTAAGAGCAGGAGAGGCAACCTGGAGAGGCTGAAACAGTGCGAGA 3143
DB 601 ACAAAGAGAACTAAGAGCAGGAGAGGCAACCTGGAGAGGCTGAAACAGTGCGAGA 660

QY 3144 GACGTTTGAACAATGAGTCAGTAGCACAAGAGAGATGACATTTTACCTAGCATTATAACCC 3203
DB 661 GACGTTTGAACAATGAGTCAGTAGCACAAGAGAGATGACATTTTACCTAGCATTATAACCC 720

QY 3204 TGGTTGCTCTGAAGAACTGCTTCAATTTGTATATATGACTATTTACATGTAATCAAC 3263
DB 721 TGGTTGCTCTGAAGAACTGCTTCAATTTGTATATATGACTATTTACATGTAATCAAC 780

QY 3264 ATGG-GAACTTTTAGGGAA-CCTAATAAGAAATCCCAATTTTTCAGGAGTGCTGTTGCA 3321
DB 781 ATGGGAACTTTTAGGGAACTTATAGAAATCCCAATTTTTCAGGAGTGCTGTTGCTG 840

QY 3322 ATAAA 3326
DB 841 GTCAA 845

RESULT 15
BQ884082
LOCUS
DEFINITION
AGENCOURT_8119366 Lupeki dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6179939 5', mRNA sequence.
ACCESSION
BQ884082
VERSION
BQ884082.1 GI:22276090
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://imgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rsraus@rmail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium
Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://imgc.llnl.gov
Plate: LLNL13561 Row: 0 Column: 12
High quality sequence stop: 657.

FEATURES
Location/Qualifiers
1..904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179939"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupeki dorsal root ganglion"
Note: Vector: pCMV-SPORT6 (Life Technologies); Site 1:
Note: Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGGCGGCTCG-3' and
5'-GATAGTTCTAGATCCGAGCGGCGCTCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies.

BASE COUNT 252 a 265 c 240 g 147 t
ORIGIN

Query Match 23.5%; Score 791.4; DB 13; Length 904;
Best Local Similarity 96.7%; Pred. No. 3.2e-153;
Matches 862; Conservative 0; Mismatches 21; Indels 8; Gaps 5;

QY 1658 GCTCAGTGGCCATCGAGGTGGACGGCAGGGTGTAACAAGTGGCTGGGTGATGCCGCC 1717
DB 1 GCTCAGTGGCCATCGAGGTGGACGGCAGGGTGTAACAAGTGGCTGGGTGATGCCGCC 60

QY 1718 AGCCCGGAACCTCACCAAGGCGCTGCGAGGGGCGCTGAGGACCAAGATGACAAGG 1777
DB 61 AGCCCGGAACCTCACCAAGGCGCTGCGAGGGGCGCTGAGGACCAAGATGACAAGG 120

QY 1778 ATGGTGGGAGCTTCAGTGGCAGCTGGAGGCTTCCGACTACTCAGCGGCCAACCCATTA 1837
DB 121 ATGGTGGGAGCTTCAGTGGCAGCTGGAGGCTTCCGACTACTCAGCGGCCAACCCATTA 180

QY 1838 AAGTGACACATCGGTGCTACATCTAGAGAACACACAGTCCAGTGTGACCTGGACCTGT 1897
DB 181 AAGTGACACATCGGTGCTACATCTAGAGAACACACAGTCCAGTGTGACCTGGACCTGT 240

QY 1898 ACAAGTCCCTGCGAGGCTTGGAAGACCAAGCTGCACATCGACCAAGATTTGAACCC 1957
DB 241 ACAAGTCCCTGCGAGGCTTGGAAGACCAAGCTGCACATCGACCAAGATTTGAACCC 300

QY 1958 TGCAGAACAAATTAAGAACCTGAGGAGTCCGAGTCACTGAGAAAGAGCGGCCAG 2017
DB 301 TGCAGAACAAATTAAGAACCTGAGGAGTCCGAGTCACTGAGAAAGAGCGGCCAG 360

QY 2018 AAGAAATGTGCTGTCAAAATCAGTACCAACCCAGCACAAGAGCGCGCTCAAGCACA 2077
DB 361 AAGAAATGTGCTGTCAAAATCAGTACCAACCCAGCACAAGAGCGCGCTCAAGCACA 420

Search completed: February 14, 2004, 19:38:17
Job time : 6219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2004, 22:52:56 ; Search time 43 Seconds
(without alignments)
856.057 Million cell updates

Title: US-10-025-966A-6
Perfect score: 4750
Sequence: 1 MGPPSLVLCLLSATVFSLLG.....PEMKRPSSKSLGOLWEGWEG 870

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 segs. 42310858 residues

Total number of hits satisfying chosen parameters: 328717

```
Minimum DB seq length: 0
Maximum DB seq length: 20
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45%

```
Database : Issued Patents AL.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/POTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length		DB	ID	Description
		Score	Match Length			
1	4723.5	99.4	867	4	US-09-996-243-84	Sequence 84, Appl
2	3776	79.5	690	4	US-09-888-319A-36	Sequence 36, Appl
3	3179.5	66.9	871	4	US-09-773-436A-3	Sequence 3, Appl
4	3121.5	65.7	867	4	US-09-668-673B-2	Sequence 2, Appl
5	3063	64.5	1611	4	US-09-668-673B-16	Sequence 16, Appl
6	2093	44.1	455	4	US-09-668-673B-14	Sequence 14, Appl
7	1702	35.8	470	4	US-09-888-319A-35	Sequence 35, Appl
8	1610.5	33.9	360	4	US-09-810-347-4	Sequence 4, Appl
9	1594.5	33.6	372	4	US-09-810-347-2	Sequence 2, Appl
10	1584	33.3	709	4	US-09-668-673B-3	Sequence 3, Appl
11	1494	31.5	307	4	US-09-810-347-5	Sequence 5, Appl
12	1187	25.0	309	4	US-09-810-347-6	Sequence 6, Appl
13	827.5	17.4	510	1	US-08-484-493-11	Sequence 11, Appl
14	827.5	17.4	510	1	US-08-484-494-11	Sequence 11, Appl
15	827.5	17.4	510	2	US-08-345-212-11	Sequence 11, Appl
16	827.5	17.4	510	3	US-09-249-003-11	Sequence 11, Appl
17	827.5	17.4	510	4	US-09-685-844-11	Sequence 11, Appl
18	565	11.9	160	4	US-09-668-673B-7	Sequence 7, Appl
19	507	10.7	98	4	US-09-668-673B-22	Sequence 22, Appl
20	405	8.5	96	4	US-09-668-673B-20	Sequence 20, Appl
21	399.5	8.4	196	4	US-09-668-673B-18	Sequence 18, Appl
22	370.5	7.8	520	4	US-09-773-436A-10	Sequence 10, Appl
23	325	6.8	552	4	US-09-773-436A-9	Sequence 9, Appl
24	281.5	5.9	406	4	US-09-227-357-187	Sequence 187, App
25	272.5	5.7	556	1	US-08-445-586-2	Sequence 2, Appl
26	271.5	5.7	554	1	US-08-445-586-7	Sequence 7, Appl
27	238.5	5.0	583	1	US-08-445-586-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-09-996-243-84
; Sequence 84. Application US/09996243

; PAGE# NO. 64/8825
; GENERAL INFORMATION:

```

/ GENERAL INFORMATION:
/
/ APPLICANT: Aekhenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Kijavits, Ivar J.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/
/ TITLE OF INVENTION: Secrated and T.
/
/ TITLE OF INVENTION: Acids Encoding
/
/ FILE REFERENCE: P2730P1C.3
/
/ CURRENT APPLICATION NUMBER: US/09/
/ CURRENT FILING DATE: 2001-11-14
/
/ PRIOR APPLICATION NUMBER: 60/04978
/ PRIOR FILING DATE: 1997-06-16
/
/ PRIOR APPLICATION NUMBER: 60/06225
/ PRIOR FILING DATE: 1997-10-17
/
/ PRIOR APPLICATION NUMBER: 60/06518
/ PRIOR FILING DATE: 1997-11-12
/
/ PRIOR APPLICATION NUMBER: 60/06531
/ PRIOR FILING DATE: 1997-11-13
/
/ PRIOR APPLICATION NUMBER: 60/06577
/ PRIOR FILING DATE: 1997-11-24
/
/ PRIOR APPLICATION NUMBER: 60/07594
/ PRIOR FILING DATE: 1998-02-25
/
/ PRIOR APPLICATION NUMBER: 60/07891
/ PRIOR FILING DATE: 1998-03-20
/
/ PRIOR APPLICATION NUMBER: 60/08332
/ PRIOR FILING DATE: 1998-04-28

```

[illegible]

; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 99.4%; Score 4723.5; DB 4; Length 867;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

QY 1 MGPPSLVLCILSATVFSILGSSAFSLSHRLKGFQDRNRIRNIILVLTDDOVLG 60
Db 1 MGPPSLVLCILSATVFSILGSSAFSLSHRLKGFQDRNRIRNIILVLTDDOVLG 60
QY 61 MQVMNKRIRIEOGGAHFINAFVITPMCCPSRSSILTGKYVHNENTYNNENSSPSWQA 120
Db 61 MQVMNKRIRIEOGGAHFINAFVITPMCCPSRSSILTGKYVHNENTYNNENSSPSWQA 120
QY 121 QHESRTPAVLINSTGYRTAFPGKYLNEVNGSYVPPGKWKVGLLKNRFPNYTLCRGVK 180
Db 121 QHESRTPAVLINSTGYRTAFPGKYLNEVNGSYVPPGKWKVGLLKNRFPNYTLCRGVK 180
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLFP 240
QY 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVEGTGELONTYIVVTADHGYHIGFGLVKGSMPEYEDIRVPFVYVRGNVEAGCLNPHIV 360
Db 301 LVEGTGELONTYIVVTADHGYHIGFGLVKGSMPEYEDIRVPFVYVRGNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKXKRWVWDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKXKRWVWDSFLVERGKLL 420
QY 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
Db 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
QY 481 MELGGSRALSNLVPKYQGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 540
Db 481 MELGGSRALSNLVPKYQGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 540
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 600
QY 601 HRCYILENDTVQCDLDLYKSLQAWKDHKLHDHETLQNKIKNLRVGRHLKKRPEEC 660
Db 601 HRCYILENDTVQCDLDLYKSLQAWKDHKLHDHETLQNKIKNLRVGRHLKKRPEEC 660
QY 661 DCHKISYHTQHKGRKLRHSGSLHPRKGLQEKDKVLLREQKRRKKLRLQNNNTC 720
Db 661 DCHKISYHTQHKGRKLRHSGSLHPRKGLQEKDKVLLREQKRRKKLRLQNNNTC 720
QY 721 SMPGLTCTFHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFEPATGFL 780
Db 721 SMPGLTCTFHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFEPATGFL 780
QY 781 YFDLNTDPYQLMNAVNTLDRVNLQHLVQLMELRSCCKYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRVNLQHLVQLMELRSCCKYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQORRKNPEMKRPPSSKSLGOLWEGWEG 870
Db 838 QYRQORRKNPEMKRPPSSKSLGOLWEGWEG 867
  
```

RESULT 2

US-09-886-319A-36
 ; Sequence 36, Application US/09886319A
 ; Patent No. 6586185
 ; GENERAL INFORMATION:

; APPLICANT: Wolf, Eckard
 ; APPLICANT: Werner, Sabine
 ; APPLICANT: Halle, Jorn-Peter
 ; APPLICANT: Regenbogen, Johannes
 ; APPLICANT: Goppelt, Andreas
 ; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
 ; the Diagnosis or Treatment of Skin Disorders and Wound
 ; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
 ; TITLE OF INVENTION: Active Substances
 ; FILE REFERENCE: 50125/014002
 ; CURRENT APPLICATION NUMBER: US/09/886,319A
 ; CURRENT FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: US 60/222,081
 ; PRIOR FILING DATE: 2000-08-01
 ; PRIOR APPLICATION NUMBER: DE 10030149.5
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 36
 ; LENGTH: 690
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-886-319A-36

Query Match 79.5%; Score 3776; DB 4; Length 690;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLFP 240
Db 1 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLFP 60
QY 241 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 61 NASQHITPSYNAFNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 120
QY 301 LVEGTGELONTYIVVTADHGYHIGFGLVKGSMPEYEDIRVPFVYVRGNVEAGCLNPHIV 360
Db 121 LVEGTGELONTYIVVTADHGYHIGFGLVKGSMPEYEDIRVPFVYVRGNVEAGCLNPHIV 180
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKXKRWVWDSFLVERGKLL 420
Db 181 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKKXKRWVWDSFLVERGKLL 240
QY 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
Db 241 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 300
QY 481 MELGGSRALSNLVPKYQGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 540
Db 301 MELGGSRALSNLVPKYQGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVYRSRSIRSV 360
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 600
Db 361 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKDGDFSGTGGGLPDYSAANPIKVT 420
QY 601 HRCYILENDTVQCDLDLYKSLQAWKDHKLHDHETLQNKIKNLRVGRHLKKRPEEC 660
Db 421 HRCYILENDTVQCDLDLYKSLQAWKDHKLHDHETLQNKIKNLRVGRHLKKRPEEC 480
QY 661 DCHKISYHTQHKGRKLRHSGSLHPRKGLQEKDKVLLREQKRRKKLRLQNNNTC 720
Db 481 DCHKISYHTQHKGRKLRHSGSLHPRKGLQEKDKVLLREQKRRKKLRLQNNNTC 540
QY 721 SMPGLTCTFHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFEPATGFL 780
Db 541 SMPGLTCTFHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCFEPATGFL 600
QY 781 YFDLNTDPYQLMNAVNTLDRVNLQHLVQLMELRSCCKYKQCNPRTRNMDLGLKDGGSYE 840
  
```

```
Db 601 YFDLNTDTPQLMNAVNTLDRDVLNQLVQLMELRSCKGYKQCNPTTRNDJLGLKDGGSYE 660
QY 841 QYRQFORRWKPEKRPSSKSLQOLWEGWEG 870
Db 661 QYRQFORRWKPEKRPSSKSLQOLWEGWEG 690

RESULT 3
US-09-773-426A-3
; Sequence 3, Application US/09773426A
; Patent No. 6534302
; GENERAL INFORMATION:
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Williamson, Mark
; APPLICANT: Teia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el
; TITLE OF INVENTION: Human Sulfatases (A CIP Application)
; FILE REFERENCE: 35800/208398(5800-79)
; CURRENT APPLICATION NUMBER: US/09/773,426A
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 09/495,823
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 871
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-773-426A-3

Query Match 66.9%; Score 3179.5; DB 4; Length 871;
Best Local Similarity 65.0%; Pred. No. 9e-302;
Matches 579; Conservative 120; Mismatches 141; Indels 51; Gaps 12;

QY 5 SLVCLLSATVPSLGGSSAFLSHRLKGRFORDRNIRPNILVLTDDQDVELGSMQVM 64
Db 7 ALVAVLGT---ELLGSLSTVRSFRGRIQGRKNIRPNILVLTDDQDVELGSLQVM 63
QY 65 NKTRIMEOGGAHFINAFVTTMCCPSRSSILTKYVHNHNTYNNENCSGPSWQAQHS 124
Db 64 NKTRIMEOGGATFINAFVTTMCCPSRSSMLTKYVHNHNTYNNENCSGPSWQAHEP 123
QY 125 RTAVVNLNGVETAFPGKYLNEYNGSYPPGKWKVGLLKNRFPVNTLCRNGVKEKHG 184
Db 124 RTAVVNLNTGYRTAFPGKYLNEYNGSYPPGKRWGLGKNSRFYNTVCRNGIKRKHG 183
QY 185 SDYSKDYLTDLTNDVSFFRTSKMYPHRPVLWVISHAAPHGPDSPAPQYSLRFPNASQ 244
Db 184 FDYAKDYFTDLITNESINYFMSKRMYPHRPVMVISHAAPHGPDSPAPQSKLYPNASQ 243
QY 245 HTPSYNYAPNDPKHIMTYGPMKIHMEFTNMLQKLOTLMVDDSMETIYNMLVET 304
Db 244 HTPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLQKLOTLMVDDSVESLYNMLVET 303
QY 305 GELDNTYVVTADHGHIGQGLVKGKSMPEYEDIRVPYVGPVNPVAGCLNPHVLNID 364
Db 304 GELNYYIIVTADHGHIGQGLVKGKSMPEYEDIRVPYVGPVNPVAGCLNPHVLNID 363
QY 365 LAPTLIDAGLIDPDMWGSILKLDTRPNVRPHLKKGRWRDSDFLVERGKLLHHRD 424
Db 364 LAPTLIDAGLIDTFFPDVGGSKVLKLDPSKPGNFRFTNKAKIWRDTFLVERGKFLRKE 423
QY 425 NDKVDAQEEFLPKYRVKDLQRAEYQACEQLOGKWCQVEDATGKLLHKCKGPMRLG 484
Db 424 ESSKNTQQSNHLPKYREVKELCQARYQACEQPGQKWCQIEDTSKLLRIHKCKGPDLL 483
QY 485 GSR-ALSNLVPKYGGGSEACTCDSDGYKLSLAGRR-KKLFXK-----KYKASYVRSRI 537
Db 484 TVRQSTNLYVARGFHDKDKCSKRESCYRASRSQRKQFLRNQGTPKYKPRFVHTROT 543
QY 538 RSVAILVDGRVYHVLGD-----AAQPRNLTKEHWP-----APEDQDDKDG-----DFS 583
```

```
Db 544 RSLSVFEGEYIDNLEEEELQVLPRIAKRHBGKGRDLOQASSGNGRMLADSS 603
QY 584 GTGGLPDYSAANDIKVTHRYCYILENDTVQDDLDLYKSLQAWKDKHLHIDHETIETONKIK 643
Db 604 NAVGPP-----TTVRVTRKCFILPNDISHCERELYQSARAWKDKAYIDKEIEALQDKIK 658
QY 644 NLREVRGHLKCKRPEECCHKLSYHTQHKGRULKHR--GSSLHPFRKGLQEKD-KVWLLE 700
Db 659 NLREVRGHLKCKRPEECSCSKQSYNKEKGVKQKBLKSHLHPFKAAQEVDSKLOLFXE 718
QY 701 -QKRKKKRLKRLKRLQNDTCSPGLTCTHNDQNHQWQTAPFTWLTGPFCACTGANNNTYWC 759
Db 719 NNRRKKRKKRQKRGECSPGLTCTHNDNNHWQTAPFTWLTGPFCACTGANNNTYWC 778
QY 760 MRTINETHNPLFCFEPATGFLVFDLNTDPVQLMNAVNTLDRDVLNQLVQLMELRSCKGY 819
Db 779 LRVNETHNPLFCFEPATGFLVFDLNTDPVQLMNAVNTLDRDVLNQLVQLMELRSCKGY 838
QY 820 KQCNPRTRNMDLGLKDGGSYEQYRQFRRKWPMPKRPSSKSLQOLWEGWEG 870
Db 839 KQCNPRKQNLVGNKDGGSYDLHR-----GQLWDGWEG 871

RESULT 4
US-09-668-673B-2
; Sequence 2, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 867
; TYPE: PRT
; ORGANISM: Quails
US-09-668-673B-2

Query Match 65.7%; Score 3121.5; DB 4; Length 867;
Best Local Similarity 64.1%; Pred. No. 4.2e-296;
Matches 572; Conservative 117; Mismatches 145; Indels 59; Gaps 10;

QY 5 SLVCLLSATVPSLGGSSAFLSHRLKGRFORDRNIRPNILVLTDDQDVELGSMQVM 64
Db 7 ALVAVLGT---ELLTSHSSTLKLRFGRVQGRKNIRPNILVLTDDQDVELGSLQVM 63
QY 65 NKTRIMEOGGAHFINAFVTTMCCPSRSSILTKYVHNHNTYNNENCSGPSWQAQHS 124
Db 64 NKTRIMEOGGAFINAFVTTMCCPSRSSMLTKYVHNHNTYNNENCSGPSWQAHEP 123
QY 125 RTAVVNLNGVETAFPGKYLNEYNGSYPPGKWKVGLLKNRFPVNTLCRNGVKEKHG 184
Db 124 RTAVVNLNTGYRTAFPGKYLNEYNGSYPPGKRWGLVKNRFPVNTYISRNGNKEKHG 183
QY 185 SDYSKDYLTDLTNDVSFFRTSKMYPHRPVLWVISHAAPHGPDSPAPQYSLRFPNASQ 244
Db 184 FDYAKDYFTDLITNESINYFMSKRMYPHRPVMVISHAAPHGPDSPAPQSELYPNASQ 243
QY 245 HTPSYNYAPNDPKHIMTYGPMKIHMEFTNMLQKLOTLMVDDSMETIYNMLVET 304
Db 244 HTPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLQKLOTLMVDDSMERLYQMLAE 303
QY 305 GELDNTYVVTADHGHIGQGLVKGKSMPEYEDIRVPYVGPVNPVAGCLNPHVLNID 364
```


Db 304 GELENTYIYIADHGYHIGQGLVKGKSNPYDFDIRVFFIRGSPSVVPOIVLNID 363
QY 365 LAPPTILDIAGLIPADMDGKSLKLLDTERPVNRFLHKKQWVRDPSPLVERGKLLAKRD 424
Db 366 LAPPTILDIAGLTPDMDGKSLKLLDTERPVNRFLHKKQWVRDPSPLVERGKLLAKRD 423
QY 425 NDKVDAQEENFIPKQVRKDLQCRAYQACEOLGQKWCQVEDATGKLKHKCKGPRWLG 484
Db 424 EANKTQOSNQIPKRVKELCQARYQACEOLGQKWCQVEDATGKLKHKCKGPRWLG 483
QY 485 GSRALSNLV-PRYQGGSEACTCDSDGYKLSLAGRR-----KKLFKKKYKASVVRGSI 537
Db 484 AIRKTRSHSRGYSKDKCNCBGTDFRNSPTQKNOQFLRNPSSAKYXPRFVHTROT 543
QY 539 RSVALTEVDORVTHVGLD-----AAPRNLTKRH-----WGPAPEDQDDKGGGDFS 583
Db 544 RSLSEFEGEIYDINLESEELQVLKTRSIITKRNAENDKAEETDGPAGDTWADGTDVI 603
QY 584 GTGGLPDYSAAPIKVTHTCYILENDTVQCDLQKSLQAMKDKHLDHIEITLQNKIK 643
Db 604 G-----QPSVVRVTHKCFILPNDTIRCELYQSARAWDKHAYIDKEIEALQDKIK 655
QY 644 NLREVRGHLKRRPBCDCHKISYHTQHG-----RLKRGSSLSHPFRKGLQEKD-KVWL 697
Db 656 NLREVRGHLKRRPBCDCHKISYHTQHG-----RLKRGSSLSHPFRKGLQEKD-KVWL 712
QY 698 LREOKKKKRLKRLQNLQNDTCSMGLTCTHDNQHWTAPFWTLGPFCACTSANNTY 757
Db 713 FKENRRKKRGKRRKQKGGDECSLPGTCTFDNNHWTAPFWTLGPFCACTSANNTY 772
QY 758 WCMRTINETHNPLFCFATGLEYFDLNTDPYQLMNAVNTLDRLVNLQHLVQLMELRSCK 817
Db 773 WCLRTVNDTHNPLFCFATGLEYFDLNTDPYQLMNAVNTLDRLVNLQHLVQLMELRSCK 832
QY 818 GYKQCNPRTRMDLGLKGGSYEQVRQRRKWPDKPSKSGQLWEGWEG 870
Db 833 GYKQCNPRTRMDLGLKGGSYEQVRQRRKWPDKPSKSGQLWEGWEG 867

RESULT 5
US-09-668-673B-16
Sequence 16, Application US/09668673B
GENERAL INFORMATION:
APPLICANT: Emerson, Charles P
APPLICANT: Dhoot, Gurtej K
TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
CHARACTERIZATION OF SUCH PROTEINS
FILE REFERENCE: PENN-0733
CURRENT APPLICATION NUMBER: US/09/668,673B
CURRENT FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,738
PRIOR FILING DATE: 1998-09-23
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 16
LENGTH: 1611
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (819)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (840)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (844)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (852)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (858)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (865)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (875)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (878)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (881)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (888)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (896)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (907)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (910)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (915)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (927)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (943)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (945)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (948)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (954)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (959)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE
LOCATION: (971)
OTHER INFORMATION: any amino acid
FEATURE:
NAME/KEY: UNSURE

LOCATION: (974)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1018)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1046)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1080)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1089)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1102)..(1103)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1105)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1121)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1127)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1191)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1199)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1223)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1235)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1250)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1307)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1321)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1356)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1362)	OTHER INFORMATION: any amino acid
FEATURE:	
NAME/KEY: UNSURE	
LOCATION: (1382)..(1383)	OTHER INFORMATION: any amino acid

Db 656 VDSKQLQFKENRRRKKERKQRKGECSLGLCTFTHDNNHQTAPFNLGSPFAC 715
QY 750 TSANNNTYWCMTINETHNLFCEFATGFLFYDLNTDPYQLMNAVNTLDRDVLNQLHVQ 809
Db 716 TSSNNNTYWCMTINETHNLFCEFATGFLFYDLNTDPYQLMNAVNTLDRDVLNQLHVQ 775
QY 810 LMELRSCKGYKQCNPRTRNMMDLGLKDGSSYEQYQFORRWKPEMKRPSSKSLQQLWEGWE 869
Db 776 LMELRSCKGYKQCNPRTRNMMDLGLKDGSSYEQYQFORRWKPEMKRPSSKSLQQLWEGWE 817
QY 870 G 870
Db 818 G 818

RESULT 6
US-09-668-673B-14
; Sequence 14, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (142-143)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (146)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (445)
; OTHER INFORMATION: any amino acid
US-09-668-673B-14

Query Match 44.1%; Score 2093; DB 4; Length 455;
Best Local Similarity 86.1%; Pred. No. 6.6e-196;
Matches 384; Conservative 16; Mismatches 38; Indels 8; Gaps 4;

QY 431 QEENFLPKYQKVQLCORAEYQTAQCEQLGQKQWQVEDATGKLHKCKGPMRL---GGSR 487
Db 1 QEENFLPKYQKVQLCORAEYQTAQCEQLGQKQWQVEDATGKLHKCKGPMRFGGGSGSR 60
QY 488 ALSNLVPRKYVGSGSEACTCDS---GDYKLSIAGRRKLFKKYKASVYRSRSIRSVIAEV 544
Db 61 ALSNLVPRKYVGSGSEACTCDSGGGGYKVLGLAGRR-KLFPKKYKTSYARNRSIRSVIAEV 119
QY 545 DGRVYHVLGLDAAQPRNLTKRHWPAGPDQDDKDGDFSGTGLPDYSAANPIKVTHRCY 604
Db 120 DGEIYHVLGLDTPVQPRNLSKPHXXGAXEDQDDKDGSGFSFGTGLPDYSAANPIKVTHRCY 179
QY 605 ILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLREVRGHLKKRPEECDCCHK 664
Db 180 ILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLREVRGHLKKRPEECDCCHK 239
QY 665 ISYTHQKGRGLKRGSSSLHPRKGLQEKDKVWLLREOKRKKLKLKLRQNNDTCSMPG 724
Db 240 ISYHSQKGRGLKRGSSSLHPRKGLQEKDKVWLLRDNRNRENCATAQTAAEQR-YVQHAG 298

QY 725 LTCFTHDNOHWQTAPFNLGPFCACTSANNTYWCMTINETHNLFCEFATGFLFYEDL 784
Db 299 LTCFTHDNOHWQTAPFNLGPFCACTSANNTYWCMTINETHNLFCEFATGFLFYEDL 358
QY 785 NTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMMDLGLKDGSSYEQYRQ 844
Db 359 STDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMMDLGLRDGSSYEQYRQ 418
QY 845 FORRWKPEMKRPSSKSLQQLWEGWEG 870
Db 419 FORRWKPEMKRPSSKSLQQLWEGWEG 444

RESULT 7
US-09-886-319A-35
; Sequence 35, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Gopelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 371
; OTHER INFORMATION: Xaa= any amino acid
US-09-886-319A-35

Query Match 35.8%; Score 1702; DB 4; Length 470;
Best Local Similarity 75.3%; Pred. No. 1.3e-157;
Matches 327; Conservative 21; Mismatches 36; Indels 50; Gaps 7;

QY 300 MLVETGELDNTYIVVTADGHYHIGQFGLVKGKSMPEYEDIRVPFVVRGPNVEAGLNPHI 359
Db 1 MLVETGELDNTYILITADGHYHIGQFGLVKGKSMPEYEDIRVPFVVRGPNVEAGLNPHI 60
QY 360 VLNIIDLAP-----TILDIAGLDIPADMDGKSLIKLLDTERPVN 397
Db 61 VLNIIDLGPHTGYRTWGTGPHRGREYVSTTGLRAASEPVPLEKEAEG-----GETP-- 113
QY 398 RFLHKKKRWEDSLFVERGKLLHKRDNKYDAQENFLPKYQKVQLCORAEYQTAQCEQ 457
Db 114 -----SWREA-----NLLHKREGDKYNAQENFLPKYQKVQLCORAEYQTAQCEQ 159
QY 458 LQKQKQCVEDATGKLHKCKGPMRL---GGSRALSILVPRKYVGSGSEACTCDS---GDY 511
Db 160 LQKQKQCVEDATGKLHKCKGPMRFGGGGGSRALSILVPRKYVGSGSEACTCDSGGGSDY 219
QY 512 KLSLAGRRKLPKKYKASVYRSRSIRSVIAEVGRVTHVGLGDAAPRNLTKRHWPAG 571
Db 220 KLGLAGRR-KLFPKKYKTSYARNRSIRSVIAEVGDGEIYHVLGLDTPVQPRNLSKPHWPAG 278
QY 572 EQDDKDGDRSGTGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKHLDH 631

Db 279 EQDDKDGSGSFGTGLPDYSAPNPIKVTHTRCYILENDTVCQDLDLYKSLQAWDKHLHI 338
Qy 632 DHEIETLQNKINLRVSGHLKKRPEECDCCHKISYHTQHKGRKLRGSSSLHPRKGLQE 691
Db 339 DHEIETLQNKINLRVSGHLKKRPEECDCCHKISYHTQHKGRKLRGSSSLHPRKGLQE 399
Qy 692 KDKVLLREQREKK 705
Db 399 KDKVLLRDNRARR 412
RESULT 8
US-09-810-347-4
; Sequence 4, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: THERBOF
; FILE REFERENCE: CL001169
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Human
US-09-810-347-4

Query Match 33.9%; Score 1610.5; DB 4; Length 360;
Best Local Similarity 80.1%; Pred. No. 7.4e-149;
Matches 286; Conservative 40; Mismatches 28; Indels 3; Gaps 1;
Qy 5 SLVLCILSATVPSLLGGSAFLSHRLKGRFORDRNRIRPNIILVLTDDQDVELGSMQVM 64
Db 7 ALVLAVLGT---ELLGSLCSTVRSFRFGRIQOERKNIRPNIILVLTDDQDVELGSLQVM 63
Qy 65 NKTRIMEOGGAHFINAFVTTMCCPSRSSIITGKYVHNHNTYNNENCSFSSWQAQHEH 124
Db 64 NKTRKIMEHGGATFINAFVTTMCCPSRSSIITGKYVHNHNTYNNENCSFSSWQAQHEP 123
Qy 125 RTPAVYLNSTGYRTAFPGKLYNEVNGSVYPPGKWEVGLLKNRFRYNTLCRNGVKEKHG 184
Db 124 RTPAVYLNSTGYRTAFPGKLYNEVNGSVYPPGKWEVGLLKNRFRYNTLCRNGVKEKHG 183
Qy 185 SDYSKDYLTDLITNDSVSPRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSRLFPNASQ 244
Db 184 FDYAKDYFTDLITNESINFKMSKMYPHRPVLMVISHAAPHGPDSPAPQYSRLFPNASQ 243
Qy 245 HTTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304
Db 244 HTTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303
Qy 305 GELDNTYIVYTADGHYHIGQFGLVKGKMPYDFDIRVPPYVGRPNVEAGCLNPHIVL 361
Db 304 GELENTYIIVYTADGHYHIGQFGLVKGKMPYDFDIRVPPYVGRPNVEAGCLNPHIVL 360

RESULT 9
US-09-810-347-2
; Sequence 2, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: THERBOF
; FILE REFERENCE: CL001169
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Human
US-09-810-347-2
Query Match 33.6%; Score 1594.5; DB 4; Length 372;
Best Local Similarity 80.3%; Pred. No. 2.9e-147;
Matches 282; Conservative 40; Mismatches 26; Indels 3; Gaps 1;
Qy 5 SLVLCILSATVPSLLGGSAFLSHRLKGRFORDRNRIRPNIILVLTDDQDVELGSMQVM 64
Db 7 ALVLAVLGT---ELLGSLCSTVRSFRFGRIQOERKNIRPNIILVLTDDQDVELGSLQVM 63
Qy 65 NKTRIMEOGGAHFINAFVTTMCCPSRSSIITGKYVHNHNTYNNENCSFSSWQAQHEH 124
Db 64 NKTRKIMEHGGATFINAFVTTMCCPSRSSIITGKYVHNHNTYNNENCSFSSWQAQHEP 123
Qy 125 RTPAVYLNSTGYRTAFPGKLYNEVNGSVYPPGKWEVGLLKNRFRYNTLCRNGVKEKHG 184
Db 124 RTPAVYLNSTGYRTAFPGKLYNEVNGSVYPPGKWEVGLLKNRFRYNTLCRNGVKEKHG 183
Qy 185 SDYSKDYLTDLITNDSVSPRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSRLFPNASQ 244
Db 184 FDYAKDYFTDLITNESINFKMSKMYPHRPVLMVISHAAPHGPDSPAPQYSRLFPNASQ 243
Qy 245 HTTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304
Db 244 HTTPSYNAPNDKHWIMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303
Qy 305 GELDNTYIVYTADGHYHIGQFGLVKGKMPYDFDIRVPPYVGRPNVEAGCL 355
Db 304 GELENTYIIVYTADGHYHIGQFGLVKGKMPYDFDIRVPPYVGRPNVEAGCL 354

RESULT 10
US-09-668-673B-3
; Sequence 3, Application US/09668673B
; Patent No. 6562356
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-668-673B-3

Query Match 33.3%; Score 1584; DB 4; Length 709;
Best Local Similarity 40.1%; Pred. No. 8.9e-146;
Matches 316; Conservative 118; Mismatches 197; Indels 158; Gaps 17;
Qy 45 NIILVLTDDQDVELGSMQVMNKTTRIMEOGGAHFINAFVTTMCCPSRSSIITGKYVHNH 104
Db 36 NVILITDDQDIELGSMDFMFKTSQIMKRGTEFTSGYVTTTICCPSSRTILTGLVYVNH 95
Qy 105 NTVNNENCSFSSWQAQHEHRTPAVYLNSTGYRTAFPGKLYNEVNGSVYPPGKWEVGLL 164
Db 96 HVHTNNQNTGVEWRKVHEKKSIGVYIQAGYKATGLYKLYNEIDGSIYPPGDEHAIIV 155
Qy 165 KNSRFNYNTLCRNGVKEKHGSDYSKDYLTDLITNDSVSPRTSKMYPHRPVLMVISHAA 224

US-08-484-493-11
; Sequence 11, Application US/08484493
; Patent No. 5728381
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,493
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-493-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;
Best Local Similarity 27.5%; Pred. No. 5.6e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RPNILVLDQDELGSMQVANKTRIMEQGAFFINAFVTPMCCPSRSILTKGVH 102
DB 3 RPNVLLTDDQDELVIGMTPLKTKALIGEMGTFFSSAYVPSALCCPSRSILTKRYPH 62
QY 103 NHNTYTN--NENCSPPSWAQHESRTFAVVLNS-TGYR--TAPFGKYLNEYNG-----S 151
DB 63 NHVAVNTLEGNCSSKSWQKIQEPNTPAILRSMQGYQTFPPAGKYLNEYGAPDAGGLE 122
QY 152 YVPPGKWEVGLLKNRFFNYTLCRNGVEKGGSDYSKDYLTDLTINDSVSP--FRSKK 209
DB 123 HVLPLGWSYVALEKNGKYNITLTSINGKARKHGENYSVDYLDVLANSYLDPLDYKSNEE 182
QY 210 MYPHRPLVAVISHAAPHGEDSAPOYSRLFPNASQHI TPSYVAPNPDGHIMRYT-GPM 268
DB 183 -----PFFMWIATPAPSPWTAAPQYQKAFQNVFAPRNKFN-INGTNKHWLIROAKTPM 236
QY 269 KPIHMEFTNNLQKRIQLTMSVDDSMETIYNNMLVETGELDNTYIVYTADHGYHICQFGLV 328
DB 237 TNSSIQFLDNAPKRWQTLTLLSVDDLVEKLVKRLFTGELNNTYIFYTSNGYHTGQFSLP 296

QY 329 GKSMPEYEFDIRVPFYRGPNVEAGCLNPHIVLNIIDLAPTILDIAGLDI-PADMDGKSIL 387
DB 297 IDKRQLYEFDIKVPLLVRGPGIKPNQTSKMLVANIDLGP-ILDIAGYDLNKTQMDGMSLL 355
QY 388 KLLDTERPVNRPHLKKOMRVWSDSLIVERGKLLHKKRDNDKVDAAQENFLPKYQRVKDLCC 447
DB 356 PIL---RGASNL---TWRSDDLVE-----YQEGE----- 373
QY 448 RABYQTACEQLGQKQWQCVEDATGKLKHKCKGPMRLGGSRALSNLVPPKYVGQSEACTCD 507
DB 374 -----YQEGE----- 378
QY 508 SGDYKLSLAGRRKKLKKYKASYVRSIRSVALEVDGRVTHVGLGDAQAQPNLTKRHW 567
DB 379 -----RNYT----- 382
QY 568 PGAPEDQDDKGGDFSGTGLPDYSAANPIKVTHRCYILENDTVQCCLDLYKSLQAKDH 627
DB 383 ----- 382
QY 628 KLHIDHEITLQNKIKNLREVRGHLKKRPEBCDCHKISYHTQHKGRCLKHRRGSSLHPRK 687
DB 383 -----DPTCPSL- 390
QY 688 GLOEKDKVLLREQEKKKLKLKELQNNDTCSMGLT-CETHDNQHWQTAPFTLGP 746
DB 391 -----PGVSQCF-----PD 399
QY 747 CACTSANNNTYCMRTINETHNLFCEP--ATGFLYEDLNTDPYQLMNAVNTLDRDVLN 804
DB 400 CVCEDAYNNTYACVRTMSALWNLQVCEFDQBFVEVYVNLTAADPDQITNIAKTIDPELLG 459
QY 805 QLVHQLMELRSCKG 818
DB 460 KMYRLMMLQSCSG 473

RESULT 14
; Sequence 11, Application US/08484494
; Patent No. 5798239
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,494
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:

/ NAME: DiGiglio, Frank S
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 8416Z
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 516-742-4343
/ TELEFAX: 516-742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 510 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-484-494-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;
Best Local Similarity 27.5%; Pred. No. 5.6e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RNNILVLTDDQDVBLGSMQVWVKTRRMEQGAHFINAFVTPMCCPSRSILTKYVH 102
DB 3 RNNVLLTDDQDVBLGSMQVWVKTRRMEQGAHFINAFVTPMCCPSRSILTKYVH 62
QY 103 NNTYTN--NENCSPSPWQAQHSRTFAVLNS--TGVR--TAFPGKYLNEYNG-----S 151
DB 63 NHHVNVNTEGNCSSKSWQKIOEPNTFFAILRSMQGYOTFFAGKYLNEYGAPDAGGLE 122
QY 152 YPPGKWEVGLLKNSRFNYTLCNGVKRKGSDYSKDYLTDLTNDVSVP--FRTSKK 209
DB 123 HVPGLGWSYVALEKNSKYNYTLSTNGKARKHGENYSVDYLTDLVANSFLDYKSNEE 182
QY 210 MYPHRPVLMVISHAAPHGPDSDAPQYSRLFPNASQHIPTPSVNYAPNPKWIMRYT-GPM 268
DB 183 -----PFFMMIATPAPHPGPTWTAPOYQAFQVFAFRNKNFN--IHGTNKGWLIQAQTPM 236
QY 269 KPIHMEFTNMLQRKQLQTLMSVDDSMETIYNNMLVETGBELDNTIYVYTDHGYHIGQGLV 328
DB 237 TNSSIQFLDNAFRKRWQTLTSDVLLVEKLVKLEFTGELNNTIYFYTSNGYHTGQSLP 296
QY 329 KGSNPFYEDIRVPYVRGPNVEACLNPHIVNLIDLAPTILDIAGLDI--PADMDGKSIL 387
DB 297 IDKRLQLEYEDIKVPLLVVRGPGIKPQTSKMLVANIDLGP--ILDIAGYDLNKTQMDGSL 355
QY 388 KLLDTERPVNRPHLKKMKRVRWDSFLVERGKLLHCRDNDKVDQAEENFLPYQYRVKDLQ 447
DB 356 PIL---RGASNL-----THRSVLYE----- 373
QY 448 RAEYQACEQLGQKQCVEDATGKLLKHCKRPMELGSRALSNLVPKYQOGSEACTCD 507
DB 374 -----YQCEG----- 378
QY 508 SGDYKLSLAGREKKLFKKYKASYRSIRSVAIEVDGRVYHVLGDAQAPNLTKRHW 567
DB 379 -----RNV----- 382
QY 568 PGAPEDQDDKGGDFSGTGGLPDYGAANPIKVTHRCYILENDTVQCOLDLYKSLQAWKH 627
DB 383 ----- 382
QY 628 KLIHIDHEITLQNKIKNLREVRGHLKKRPEBCDCHKISYTHQKGRUKRKGSSLHFRK 687
DB 383 -----DPTCPSL----- 390
QY 688 GLQEKDKVLLREOKRKKLKLRLQNNDTCSMPGLT-CFTDNOHQWQAPFWTLGPF 746
DB 391 -----PGVSQCF----- 399
QY 747 CACTSANNNTYCMRTINETHNPLCFE--ATGFLEYFDLNTDPQLMNVANTLDRVNL 804
DB 400 CVCEDAYNTYACVRTSMALMQLQCEFDQDFVEFVYVNLTAQPDQITNIAKTIDPBLG 459
QY 805 QLHVQLMELRSCGK 818

DB 460 KXNVLMLQLQSCG 473

RESULT 15
US-08-345-212-11
/ Sequence 11, Application US/08345212
/ Patent No. 5932211
/ GENERAL INFORMATION:
/ APPLICANT: Wilson, Peter J
/ APPLICANT: Morris, Charles P
/ APPLICANT: Anson, Donald S
/ APPLICANT: Occhiodoro, Teresa
/ APPLICANT: Bielicki, Julie
/ APPLICANT: Clements, Peter R
/ APPLICANT: Hopwood, John J
/ TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
/ TITLE OF INVENTION: IDURONATE 2-SULFATASE
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Scully, Scott, Murphy & Presser
/ STREET: 400 Garden City Plaza
/ CITY: Garden City
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 11530
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/345,212
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 991,973
/ FILING DATE: 17-DEC-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DiGiglio, Frank S
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 8416Z
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 516-742-4343
/ TELEFAX: 516-742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 510 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-345-212-11

Query Match 17.4%; Score 827.5; DB 2; Length 510;
Best Local Similarity 27.5%; Pred. No. 5.6e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

QY 43 RNNILVLTDDQDVBLGSMQVWVKTRRMEQGAHFINAFVTPMCCPSRSILTKYVH 102
DB 3 RNNVLLTDDQDVBLGSMQVWVKTRRMEQGAHFINAFVTPMCCPSRSILTKYVH 62
QY 103 NNTYTN--NENCSPSPWQAQHSRTFAVLNS--TGVR--TAFPGKYLNEYNG-----S 151
DB 63 NHHVNVNTEGNCSSKSWQKIOEPNTFFAILRSMQGYOTFFAGKYLNEYGAPDAGGLE 122
QY 152 YPPGKWEVGLLKNSRFNYTLCNGVKRKGSDYSKDYLTDLTNDVSVP--FRTSKK 209
DB 123 HVPGLGWSYVALEKNSKYNYTLSTNGKARKHGENYSVDYLTDLVANSFLDYKSNEE 182
QY 210 MYPHRPVLMVISHAAPHGPDSDAPQYSRLFPNASQHIPTPSVNYAPNPKWIMRYT-GPM 268

Db 183 -----PFEMMIATPAPHPWTAAPQOKAFQNFVAFPRNKFN- IHGTHKHLIRQAKTEM 236
QY 269 KPIHMEETNMLQKRLQTLASVDDSMETIYNMLVETGELDNTYIVYTAHGYHIGQGLV 328
Db 237 TNSIQFLDNFRKRWGLLVDDLVKLVKRLFTGELNNTYIFTSDNGYHTGQFSLP 296
QY 329 KGKSMPEFDIRVFPYVRGPNVEAGCLNPHIVLNIDLAPTILDLIAGLDI-PADMDGKSIL 387
Db 297 IDKROLYEFDIKVELLVGPGIKPNQTSKMLVANIDLGP-ILDIAGYDLNKTQMDGMSLL 355
QY 388 KLLDTERPVNRFHLKKQKRVWRDSFLVERGKLLHKRDNKVDQAQENFLPKYQRYKDLQ 447
Db 356 PIL---RGASNL-----TWRSDDLVE----- 373
QY 448 RABYQTACEQLGQKWQCVEDATGKLLHKCKGPMRLGGSRALSNLVPKYGOGSEACTCD 507
Db 374 -----YQGES----- 378
QY 508 SGDYKLSLAGRRKKLFKKYKASVYRSRSIRSAIEVDGRVYHVGLDAAQPRNLTKRW 567
Db 379 -----RNV----- 382
QY 568 PGAPEDQDDXGDFSGTGLPDYSAANPIKVTWRCVILENDTVQCDLCLKSLQAWKDH 627
Db 383 ----- 382
QY 628 KLHIDHEIETLQNKIKNLREVGRHLKKRPEECCHKISYHTQHKGRLKHGSSLHPFRK 687
Db 383 -----DPTCPSL----- 390
QY 688 GLQEKDKVLLRQKRRKKLRLKRLQNNDTCSMEGLT-CFTHDNQHWQTAPFWTLGPF 746
Db 391 -----PGVSQCF-----PD 399
QY 747 CACTSANNNTYWCMTINETHNLFCEP--ATGFLBYFDINTDPYQLMNAVNTLDRDLN 804
Db 400 CVCEDAYNTYACVRTMSALWNLQYCEFFDDQEVFEVYNLTADPDQITNIAKTIDPELLG 459
QY 805 QLVHQLMEIRSCG 818
Db 460 KMYRLMMLOSCG 473

Search completed: February 14, 2004, 22:59:55
Job time : 47 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1584	33.3	709	2	T16584	hypothetical prote
2	858.5	18.1	552	1	KJHUJG	N-acetylglucosamin
3	452.5	9.5	649	2	S43229	arylsulfatase (EC
4	448.5	9.4	639	2	T46577	arylsulfatase (EC
5	341.5	7.2	646	1	KJXM	arylsulfatase (EC
6	302.5	6.4	497	2	G65169	probable sulfatase
7	301.5	6.3	497	2	C91206	probable sulfatase
8	301.5	6.3	497	2	E86052	probable sulfatase
9	290	6.1	535	2	AF0103	probable sulfatase
10	251	5.3	571	2	G90891	probable sulfatase
11	251	5.3	571	2	E64303	arylsulfatase homo
12	250.5	5.3	557	2	AF0370	probable sulfatase
13	250	5.3	571	2	B85726	probable sulfatase
14	241.5	5.1	517	2	AB0102	probable sulfatase
15	238.5	5.0	583	1	KJHUAC	steryl-sulfatase (
16	238	5.0	551	2	S07089	arylsulfatase (EC
17	238	5.0	557	2	A37362	arylsulfatase (EC
18	234.5	4.9	550	1	KJHUID	iduronate-2-sulfat
19	233	4.9	497	2	AF0506	probable secreted
20	232	4.9	551	2	S01793	arylsulfatase (EC
21	228	4.8	538	2	F83354	probable sulfatase
22	226	4.8	563	2	A47153	iduronate-2-sulfat
23	224	4.7	465	2	F70837	probable sulfatase
24	223.5	4.7	479	2	C75099	hypothetical prote
25	222	4.7	503	2	E83642	choline sulfatase
26	221	4.7	514	2	T44602	phosphonate monoos
27	219	4.6	787	2	B70643	probable sulfatase
28	218.5	4.6	522	1	KJHUG6	N-acetylglactosam
29	211.5	4.5	533	1	KJHUAB	N-acetylglactosam

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

460 QKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRR 519

A;Title: The evolutionary conservation of a novel protein modification, the conversion of
 A;Reference number: S68892; MUID:96283826; PMID:8681943
 A;Accession: S68892
 A;Molecule type: protein
 A;Residues: 64-71,'S',73-76 <SEL>
 C;Function:
 A;Description: (EC 3.1.6.1) [validated, MUID:94222055]; hydrolyzes sulfate esters in res
 A;Note: required for mineralization of sulfate
 A;Note: remarkably insensitive towards detergents like dodecyl sulfate
 C;Superfamily: plant sulfatase
 C;Keywords: glycoprotein; periplasmic space; sulfuric ester hydrolase
 F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-649/Product: arylsulfatase #status experimental <NAT>
 F;72/Modified site: 3-oxoalanine (Cys) #status experimental

Query Match 9.5%; Score 452.5; DB 2; Length 649;
 Best Local Similarity 27.4%; Pred. No. 1.2e-24;
 Matches 149; Conservative 82; Mismatches 170; Indels 143; Gaps 24;

QY 6 LVCLLSATVFLGCGSAFLSHRLKGRFQDRDRNRIRNIIILVLTDDQDVELGSM--QV 63
 Db VALCLLG---FAALTAAA---HQ-----RPNFVIFTDDQDQIGNSTPTRY 47
 QY 64 MNKTRIMEQGAHFINAFVTPMCCPSRSSILTKYVHNHNTYNNENCSSP-----SW 118
 Db 48 QPKLHEHRYPGIELKNYFVITPVCCPSRTNLWRQPSHN---TNFTDVLGPHGYAKW 103
 QY 119 QAOHESRTF-AVYLNSTGYRTAFPGKYLNEY---NGSVYPPGKWEVGLLKNSRF--YNYT 173
 Db 104 KSLGIDKSYLPWMLQNLGNTYYGKFLVDYVSNGYQNPVAGTDDIDALVTPYTFDYNP 163
 QY 174 -LCRNGVKEKGGSDYSDYLTDLITNDSVSPFRSKMYPHRPLVLMVISHAAPHG----- 232
 Db 164 GFSRNGATP---NIYPGFYSTVDIADKVAQIKTA--VAACKPFYAQISPIAPH----- 212
 QY 233 POYSRLFNASCHITTPSYNAPNP-DKHWIMRYGPM-----KPIHMEF 275
 Db 213 -TSTQIYFDPVANATKTFYPPIPAPRIHWELFSDATLPEGTSKHNLYEADVDRPAMIRA 271
 QY 276 TNLQO-----RKRLOTMSVDDSMETIYNMLVETGELONTYIVVTADGHYHIG 324
 Db 272 LPLAQQNRVLEEVYRLRLSLASVDELIRVATVQEAQVLTNTLITSADNGYHVT 331
 QY 325 FGLVKGKSMPEFDIRVPFYVRGPNVEAG----CLNPHIVLINIDLAPTILDIAG----- 374
 Db 332 HRFAGKAGTAYDEDLRVPLIRGPGIRASHSDKPKANSKVLHVDFAPTILTLACAGQVQ 391
 QY 375 -----LDIPADMDGKSLKLLDTERPVNRFLKKQVWRDSFLVE-----RGKLLH 421
 Db 392 DKALDGTPLGLYANDGN---LLADYPRPNH-----RNQFQGEFWGWSDEVLH 438
 QY 422 KRDNQKVDQAENFLPKYQ-----RVKDLQRAEYQTAQEQGKQWCQVEDATGKLL 474
 Db 439 -----HIPRYTNNSKAVRYD-----EDNQQAQKLVISCTNEREL 474
 QY 475 HKCK 478
 Db 475 YDLK 478

RESULT 4
 T46577
 arylsulfatase (EC 3.1.6.1) [validated] - Neurospora crassa
 C;Species: Neurospora crassa
 C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 04-Mar-2000
 C;Accession: T46577
 R;Baker, D.L.; Paletta, J.V.
 submitted to the EMBL Data Library, February 1997
 A;Description: Molecular characterization of the arylsulfatase gene of Neurospora crassa
 A;Reference number: Z23090
 A;Accession: T46577
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA

A;Residues: 1-639 <BAK>
 A;Cross-references: EMBL:U89492; PIDN:AAC02716.1
 A;Experimental source: strain wild type 74-OR23-1A
 C;Genetics:
 A;Gene: ars-1
 A;Map position: 7
 A;Introns: 115/1, 518/3
 C;Function:
 A;Description: EC 3.1.6.1 [validated, MUID:89384589]
 A;Note: genes are expressed under conditions of sulfur limitation and are under coordin
 C;Superfamily: plant sulfatase
 C;Keywords: sulfuric ester hydrolase

Query Match 9.4%; Score 448.5; DB 2; Length 639;
 Best Local Similarity 27.5%; Pred. No. 2.3e-24;
 Matches 137; Conservative 73; Mismatches 212; Indels 77; Gaps 14;

QY 13 ATVFSLLGSSAFLSHRLKGRFQDRDRNRIRNIIILVLTDDQDVELGSMQVKNKRRIME 72
 Db 15 STAFVDSQKSLGFEHKPKGTTTEGKK-SNIVFILTDDQDLHLQSLDYLPLKKYLA 73
 QY 73 QGAHPINAFVTPMCCPSRSSILTKYVHNHNTYNNENCSSPSWQAQHSRTF-AVY 130
 Db 74 DEGTYYKRYCTTAICCPARVSLWTGQAHNTNVTDSVPPYGGYKPFISQGFNEAYLPW 133
 QY 131 LKSTGYRTAFPGKYLNEY---EYNGSVYPPGKWEVGLLK--NSRFYNYTLCRNGVKEKH 183
 Db 134 LQKAGDITTYTKLNFMAITVDNIDSPYI--AGWNGSDFLLDPYTYLYLNATFORN---RDP 189
 QY 184 GSDYSKDYLTDLITNDSVSPFRSKMYPHRPLVLMVISHAAPHG----- 227
 Db 190 PISYEQYSVDVLAEKAYGFLDEAAKNVHNRPFFLGIAPIAPHSNVEPGFPSSSSSSSS 249
 QY 228 -----PEDSAPQYSRLFPNASCHITTPSYNAPNP-DKHWIMR 263
 Db 250 DSATLHRRPTNEHDDIEKSVSFTPIPAARHAHLFPDVIPTPHFNFSRSGASVSIAR 309
 QY 264 YTGPKMPPHMEFTNMLQKRLQTLMSVDDSMETIYNMLVETGELONTYIVVTADGHYHIG 323
 Db 310 LP-HQSAENVAENDFYRQLRALESVDLAVAGVVARLERHGLLENTYVFTADNGYHIG 368
 QY 324 QGLVKGKSMPEFDIRVPFYVRGPNVEAGCLNPHIVLINIDLAPTILDIAG----- 375
 Db 369 QHRLQPGKECGFEEDINVPLIVRGPGVKKEGVAEVVTSHTVDIAPTVLKLAGALGDKESKE 428
 QY 376 ---DIPADMDGKSI---LKLLDTERPVNRFLKKQVWRDSFLVERGKLLHKNRDNKVD 429
 Db 429 EEEEREYGLDGEAIPVTAEEDLAQAKAKGRHHEHVTVEYW--GFAVSEGRVPEK-----D 480
 QY 430 AQEENFLPKYQVVKDLQCR 448
 Db 481 ATRYTNNTYKAVKILGER 499

RESULT 5
 KJKN
 arylsulfatase (EC 3.1.6.1) precursor - Chlamydomonas reinhardtii
 N;Alternate names: sulfatase
 C;Species: Chlamydomonas reinhardtii
 C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 24-Sep-1999
 C;Accession: JQ0310
 R;de Hostos, E.L.; Schilling, J.; Grossman, A.R.
 Mol. Gen. Genet. 218, 229-239, 1989
 A;Title: Structure and expression of the gene encoding the periplasmic arylsulfatase of
 A;Reference number: JQ0310; MUID:89384447; PMID:2476654
 A;Accession: JQ0310
 A;Molecule type: mRNA
 A;Residues: 1-646 <DEH>
 A;Cross-references: GB:X16180; NID:g18118; PIDN:CAA34302.1; PID:g18119; GB:X52304; NID:g
 A;Experimental source: strain cw15mt+
 A;Note: part of this sequence, including the amino end of the mature protein, was determ
 C;Comment: This enzyme is commonly produced by soil microorganisms and plays an importar
 C;Superfamily: plant sulfatase

C;Keywords: glycoprotein; sulfuric ester hydrolase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-646/Product: arylsulfatase #status predicted <MAT>
F;42,443,527/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;73/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 7.2%; Score 341.5; DB 1; Length 646;
Best Local Similarity 24.6%; Pred. No. 1.3e-16;
Matches 124; Conservative 72; Mismatches 180; Indels 129; Gaps 18;

QY 43 RPNILVLTDQDVELGSM-QVMNTRIMEOGGAHFINAFVTTMCCPSRSSILTCKY 100
DB 26 KPNFVIFTDQDAIGNSTHYPMLSHKIRYPGVVELSQYVTTVPVCCPSRTNCAAS- 84
QY 101 VHNHNTYNNNC-----SFSWQAQHSR-----TPAVYLNSTGYRTAFF 141
DB 85 --SPTTPTSPACCLPTVAGPSGRWASTPTCRGSRTRKAITPTTWASSLWTTPTSATTSR 142
QY 142 GKYLNENYNGSYVPPGKEWGLLKNRFFNYTLCRNGVKEKHSYDYSKDYLTDLTNDVS 201
DB 143 CRGLGRYCPVTPYTFD-----YNTLRQNGATP-----NYPGYSTDVIRDKGV 188
QY 202 SFFRTSKMYPHRLPVLVISHAAPHGSDAPQYSLFPNASCHITPSPNYAPNP-DKHW 260
DB 189 AQIKSA--VAAGKFFVAQISPIAPH-----TSTQISTNPATGVTRSYFFPIPAPEHW 239
QY 261 IM-----RVTGPKPFIHMEFTNMLQKRLQTLMSVDD 292
DB 240 QLESDANLPGSQOEPFLGRGRERQAEQLDRPAAGPAEQPHLPGGDLPAP--EVAGPVDE 297
QY 293 SMETIYNMLVETGEIDNTYIVTADHGHIQGLV-KGKSMPEYEDIRVPFVVRGNVEA 352
DB 298 LIQVVKTLDEAGVLNTYIISADNGYHVGARFGAGTKTYEEDLRVPFLIRGPIKA 357
QY 353 G-----CLNPHVNLIDAPTILDIAG-----LIPADMCKSLTKLLDTER 394
DB 358 SKSDXPQNSKVGVLHVDPAFTILSLAGASHLLGDKLDTGFLGYANDDG-----TLPSDY 412
QY 395 PVNRFLHKKMRV---WRDSFLVERGKLLHKKRDNKDVAQENFLPKYQVRVKDLCQRAE 450
DB 413 PREQHQRFQGFEGFWGWDGL-----QNLRSQNNTW-KVRYTD----- 453
QY 451 YQFACBQLGQKWCQVEDATGKULH 475
DB 454 -----ESSKQGWKLIAQCTNERELY 473

RESULT 6
G65169
N;Alternative names: hypothetical 57.3K protein ilvo-ibpB intergenic region
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: G65169
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
A.; Rose, D.J.; Mau, B.; Shao, Y.
S;Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65169
A;Status: preliminary
A;Molecule type: DNA
A;Title: Complete genome sequence of Escherichia coli K-12.
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: EC84619

Query Match 6.3%; Score 301.5; DB 2; Length 497;
Best Local Similarity 25.5%; Pred. No. 7.1e-14;
Matches 105; Conservative 73; Mismatches 153; Indels 81; Gaps 21;

QY 43 RPNILVLTDQDVELGSMQVMNTRIMEOGGAHFINAFVTTMCCPSRSSILTCKY 97
DB 3 RPNFLFMTDTPQAPNVMGVCYSGKPLN-TQIDSLAAEGIRFNSAYTCSPVCTPARAGLT 61
QY 98 GKVVHNTYNN-----ENCSSPSWQAQHSRTPFAYLNSTGYRTAFFGKYL---NEYNG 150
DB 62 GIYANQSGPTNNVAPGKNIS-----TMGRYFKDAGYHTCYIGKWLHDGHDYFG 110
QY 151 S-YVPPGKW---EMVGLLKNRFFNYTLCRNGVKEKHSYDYSKDYLTDL-----ITN 198
DB 111 TGCCPPEWDADYDFDAGNYLSELTEKISLWRNGLNSV--EDLOANHIDETFTWAHRISN 168
QY 199 DSVSFFRTSKMYPHRLPVLVISHAAPHG-----EDSAPQYSLFPNASCHITPSP 250
DB 169 RAVDFLQOPAR--ADEPFLMVSYDEPHHPTCPVEYLEKYADFFYELGEKAQDDL--- 222
QY 251 NYAPNPKH--WIMRYTGP-----KPIHMEFTNMLQKRLQTLMSVDDSMETIYNML 301
DB 223 --ANKPHEHLWAQAMPSPVGGDLVHHPLYFACNDF-----VDDQIGRVINAL 269
QY 302 VETGEIDNTYIVTADHGHIQGLV-KGKSMPEYEDIRVPFVVRGNVEAGCLNPHIV 360
DB 270 TPE-QRENTWVIITSDHGEEMGAHKLISGAAM-YDDITRIPLIIRSPQERRQVDP-V 326
QY 361 LINDLAPTILDIAGLIPADMCKSLTKLLDTERPVNRFLHKKMRVWRDSP 412
DB 327 SHIDLPTMALADIKEPILPGENILAVKEPGRVNVFEF---RYEIHDSF 375

RESULT 7
C91206
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: C91206
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
S;DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen-
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C91206
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-497 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA038042.1; PID:G13364094; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: EC84619

Query Match 6.3%; Score 301.5; DB 2; Length 497;
Best Local Similarity 25.5%; Pred. No. 7.1e-14;
Matches 105; Conservative 73; Mismatches 153; Indels 81; Gaps 21;

QY 43 RPNILVLTDQDVELGSMQVMNTRIMEOGGAHFINAFVTTMCCPSRSSILTCKY 97
DB 3 RPNFLFMTDTPQAPNVMGVCYSGKPLN-TQIDSLAAEGIRFNSAYTCSPVCTPARAGLT 61
QY 98 GKVVHNTYNN-----ENCSSPSWQAQHSRTPFAYLNSTGYRTAFFGKYL---NEYNG 150
DB 62 GIYANQSGPTNNVAPGKNIS-----TMGRYFKDAGYHTCYIGKWLHDGHDYFG 110
QY 151 S-YVPPGKW---EMVGLLKNRFFNYTLCRNGVKEKHSYDYSKDYLTDL-----ITN 198
DB 111 TGCCPPEWDADYDFDAGNYLSELTEKISLWRNGLNSV--EDLOANHIDETFTWAHRISN 168
QY 199 DSVSFFRTSKMYPHRLPVLVISHAAPHG-----EDSAPQYSLFPNASCHITPSP 250
DB 169 RAVDFLQOPAR--ADEPFLMVSYDEPHHPTCPVEYLEKYADFFYELGEKAQDDL--- 222
QY 251 NYAPNPKH--WIMRYTGP-----KPIHMEFTNMLQKRLQTLMSVDDSMETIYNML 301
DB 223 --ANKPHEHLWAQAMPSPVGGDLVHHPLYFACNDF-----VDDQIGRVINAL 269

```

QY 302 VETGELDNITYIVYTADHGHYHGQGLV-KGKSMPEYEFDIRVPFYVRGNVEAGCLNPHIV 360
Db 270 -KPEQRENTWIVYSDHGEMGAHKLSKGAAM-YDDITRIPLIIRSPQGERQVDTP-V 326
QY 361 LNIIDLAPTILDIAGLIDPADMDGKSILKLLDTERPVPNRFHLKKQVRVWRDSF 412
Db 227 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
227 SHIULLPTMALADIERPEILPFGNIIIAVKPRGVMVEFN--RYEIHDSF 375
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 8
E86052
probable sulfatase YidJ [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: E86052
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E86052
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1:497 <STO>
A:Cross-references: GS:AE005174; NID:g12518517; PIDN:AAGS8881.1; GSPDB:GN00145; UWGP:Z51
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: YidJ

```

[illegible]

RESULT 9
AF0103
probable sulfatase YF0842 [imported] - *Yersinia pestis* (strain CO92)
C/Species: *Yersinia pestis*
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C/Accession: AF0103
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Farraga, A.W.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0103
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-535 <KUR>
A:Cross-references: GB:AL50842; PIDN:CAC89689.1; PID:g15978916; GSFDB:GN00175
C:Genetics:
A:Gene: YPO0842

[illegible]

```

RESULT 10
G90891      probable sulfatase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 050995
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 24-Aug-2001
C/Accession: G90891
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
  Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
  DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: G90891
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-571 >RAY>
A/Cross-references: GB:BA000007; PIDN:BA035526.1; PID:G13361569; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:

```

A:Gene: ECs2103

C:Superfamily: animal sulfatase

```
Query Match          5.3%; Score 251; DB 2; Length 571;
Best Local Similarity 22.7%; Pred. No. 3.9e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

QY 18 LLGSSAFSLSHRLKGRFQDRRNI-----RPNILVLTDQ-----54
DQ 29 LASGMAFAAHAADDVKLKATKTNVAFSDFTPTTEYSTKGKRNIIIVLTMDDLGYQLPFDK 88
DQ 55 -----DVLGSMQVNMKTRRIMEQGAHFNAFVTTMCCPSRS 93
DQ 89 GSFDPKTMENREVVDYTKIGDKAIEAAQKSTPTLLSLMDEGVRFRTGYVAHGVSGPSRA 148
QY 94 SILTGKVVHNHTYNNENSCSPSWQAQHE---SRTFAVYL-NSTGYRTAFPGK-VLYNEV 148
DQ 149 AIMGTRAPARGVYSNTD-----AODGIPLTETELPELFQNHGYTYAAGKWLHLSKI 200
QY 149 NGSVYPPGKWEVGLLKNRPFY--NYTL--CRNGVKKKHGSY-----187
DQ 201 SNVPVPED-----KQTRDYHDNFTTFSAEEMQPNRGDFYFMGFHAGATAYNSPSL 252
QY 188 -----SKDYLTLTNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSS-AP-QYSGR 237
DQ 253 FKNERVPAKYISDQLTDEAIGVDRAKTL--DQPMVLAYNAPHLNDNPAPDQYQK 310
QY 238 LFPNASQHIPTSYNAPNDKHMVYTPGMPKPIHMEFTNMLQKRLQTLMSYDSDSMETI 297
DQ 311 QFNTGSQ--TADNYTA-----SVYSVDQGVKRI 336
QY 298 YNMLVETGELDNTVIYVTAADHGYHI--QGF---GLVKG-KSMPYEFDIRVPPYVGRPNVEA 352
DQ 337 LEQKKKGQVNDTILFTSNGAVIDGLPLNGAQKGYKSTYPGGTHTPMFMW-----K 392
QY 353 GCLNP-----HIVLNIDLAFTILDIAGLDIPAD--MDGKSLKLLDTER---PVRNPHLKK 403
DQ 393 GKLPQGNVYKLIISANDFVPTALDAADISIPKDLKLDGVSLLPLWLDKQKQEPHKNLTWIT 452
QY 404 KQVWRDSFLVERGKLAHK-----RQND--KYDAQENFL 436
DQ 453 SYSWFDEENIPFDWNVHKKFVRHQSDDYPHNPTEDLSQFSYTVRNNDYSLVYTVENNQL 512
QY 437 PKYQVRKDLQCARAYQACQLGQKQCV 465
DQ 513 GLY-KLTDLQCKDNLAANPQVVKMGV 540

RESULT 11
B64903
arylsulfatase homolog b1498 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: B64903
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Ferna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64903
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-571 <BLAT>
A:Cross-references: GB:AE000247; GB:U00096; NID:g1787773; PIDN:AAC74571.1; PID:g1787775;
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: animal sulfatase

Query Match          5.3%; Score 251; DB 2; Length 571;
Best Local Similarity 22.7%; Pred. No. 3.9e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

QY 18 LLGSSAFSLSHRLKGRFQDRRNI-----RPNILVLTDQ-----54
```

```
DQ 29 LASGMAFAAHAADDVKLKATKTNVAFSDFTPTTEYSTKGKRNIIIVLTMDDLGYQLPFDK 88
QY 55 -----DVLGSMQVNMKTRRIMEQGAHFNAFVTTMCCPSRS 93
DQ 89 GSFDPKTMENREVVDYTKIGDKAIEAAQKSTPTLLSLMDEGVRFRTGYVAHGVSGPSRA 148
QY 94 SILTGKVVHNHTYNNENSCSPSWQAQHE---SRTFAVYL-NSTGYRTAFPGK-VLYNEV 148
DQ 149 AIMGTRAPARGVYSNTD-----AODGIPLTETELPELFQNHGYTYAAGKWLHLSKI 200
QY 149 NGSVYPPGKWEVGLLKNRPFY--NYTL--CRNGVKKKHGSY-----187
DQ 201 SNVPVPED-----KQTRDYHDNFTTFSAEEMQPNRGDFYFMGFHAGATAYNSPSL 252
QY 188 -----SKDYLTLTNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSS-AP-QYSGR 237
DQ 253 FKNERVPAKYISDQLTDEAIGVDRAKTL--DQPMVLAYNAPHLNDNPAPDQYQK 310
QY 238 LFPNASQHIPTSYNAPNDKHMVYTPGMPKPIHMEFTNMLQKRLQTLMSYDSDSMETI 297
DQ 311 QFNTGSQ--TADNYTA-----SVYSVDQGVKRI 336
QY 298 YNMLVETGELDNTVIYVTAADHGYHI--QGF---GLVKG-KSMPYEFDIRVPPYVGRPNVEA 352
DQ 337 LEQKKKGQVNDTILFTSNGAVIDGLPLNGAQKGYKSTYPGGTHTPMFMW-----K 392
QY 353 GCLNP-----HIVLNIDLAFTILDIAGLDIPAD--MDGKSLKLLDTER---PVRNPHLKK 403
DQ 393 GKLPQGNVYKLIISANDFVPTALDAADISIPKDLKLDGVSLLPLWLDKQKQEPHKNLTWIT 452
QY 404 KQVWRDSFLVERGKLAHK-----RQND--KYDAQENFL 436
DQ 453 SYSWFDEENIPFDWNVHKKFVRHQSDDYPHNPTEDLSQFSYTVRNNDYSLVYTVENNQL 512
QY 437 PKYQVRKDLQCARAYQACQLGQKQCV 465
DQ 513 GLY-KLTDLQCKDNLAANPQVVKMGV 540

RESULT 12
AF0370
probable sulfatase (EC 3.1.1.6.-) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF0370
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-557 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92289.1; PID:g15981000; GSPDB:GN00175
C:Genetics:
C:Superfamily: animal sulfatase
C:Keywords: sulfuric ester hydrolase

Query Match          5.3%; Score 250.5; DB 2; Length 557;
Best Local Similarity 21.3%; Pred. No. 4.1e-10;
Matches 114; Conservative 88; Mismatches 171; Indels 161; Gaps 27;

QY 10 LLSATVPSLLG---GSSAFSLSHRLKGRFQDRRNI-----RPNILVL 51
DQ 10 LLSATVPSLLG---GSSAFSLSHRLKGRFQDRRNI-----RPNILVL 51
QY 52 D-----DQDV-----ELGSMQVNMKTR-----IMEQGAHF 80
DQ 66 DDLGVQLPFDKTSFDPKSMEDRDVDYTKIGDKAIEAAKSTPTLLSLMDE-GVRLTN 124
```


81 AFVTTMCCPGRSSILTKGVYHNHTYTNENCSPPSWAQH-----BSRTFAVLNNGTY 136
:
:
:
Db GVAHGVSQSPRAAINTCKSPARFGVSYNTD-----AQNGISLETFLPELLQNGY 176
:
:
:
Qy RTAPFGK-----LNEYNSYVPGWKEWGLLKNSRPNY-----T 173
:
:
:
Db YTAALGKKHLKSKISNVPEAEQTRDYHDNFTTYSADW---QPONRGFQFMGYHAAGT 233
:
:
:
Qy LCRNGVKKHGSD--YSKYDLTDLTINDSVSFFRTSKMYPHRPVLMVISHAAPGHEDS 231
:
:
:
Db AYATASPSLFHNKERKAKGYISDQLTDBAIGVANRAKSL--DEPFMYLAIVSAHLENDN 291
:
:
:
Qy -AP-CYSLEFNASHOIHPTSYNYAPNPDKHIMRYTGPMKPIHEMTNLQRKLQTLMS 289
:
:
:
Db PAPDEYQHFWTGSQ--PADNFYA-----SVYS 317
:
:
:
Qy VDSMETIYNMLVTGELDNITYIVYTADHGVIH-QGF---GLVGK-KSMYPVEFDIRVPEY 344
:
:
:
Db VDQGVRLLLEQLKXNGQVDNIIFMTSDNGAVIDGPLPLNGNQKGYASQTFPGGTHTMP 377
:
:
:
Qy V--RGPVNVEACLNAPHIVNLDTAPTITDIAGLDIPADMCKSILKLLDTERPVNRPHLK 402
:
:
:
Db IWRKG-KIQTCNYD-KLISAMDFTALEAAEIDAPNNLDGVSLLPYLTKSKAE----P 431
:
:
:
Qy KKRVRWRDSFLVERGKLLHKRDNDKVD-----AQEENFLPKYQRVKDLQ 447
:
:
:
Db HKYLTWTYSY-----THWFDEENIFPDGYYHKFVRESNEYPKNPTEDLSQ 478
:
:
:

RESULT 13
B85726
probable sulfatase Z2210 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85726
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Groebbeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A|Reference number: A85480; MUID:21074935; PMID:11206551
A|Accession: B85726
A>Status: preliminary
A|Molecule type: DNA
A|Residues: 1-571 <STO>
A|Cross-references: GB:AE005174; NID:g12515176; PIDN:AAG56270.1; GSPDB:GN00145; UWGP:Z2210
A|Experimental source: strain O157:H7, substrain EDL933
C|Genetics:
A|Gene: Z2210
C|Superfamily: animal sulfatase

Query Match 5.3%; Score 250; DB 2; Length 571;
Best Local Similarity 22.7%; Pred. No. 4.6e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

Qy 18 LLGGSAFLSHRLKGRFORDRNI-----RPNILLVLTDDQ----- 54
Db 29 LASGMFAFAAHADDVKLKATKNVAFSPFTPTEYSTKGKPNIIIVTMDDLYGOLFDPDK 88
:
:
:
Qy 55 -----DVELGSMQVMNKTRIMEQQGAHFINAFTVTMCPGSR 93
:
:
:
Db 89 GSPDFPKTMENREVVDYTKIGIDAIAEAQAQSPTLLSLMDGEVRFVTNGVYAHGVSGPSRA 148
:
:
:
Qy 94 SILTKGVYHNHTYTNENCSPPSWAQHE---SRFAVYL-NSTGYRTAFGK-YLNEY 148
:
:
:
Db 149 AIWYGAPARFGVSYNTD-----AQGIPLUTEFLPELFQNHGYITAAVGRWLHSKI 200
:
:
:
Qy 149 NGSYVPPGWKEWGVLKNGRFY--NYTL--CRNGVKEKESGDY----- 187
:
:
:
Db 201 SNVPVED-----KQTRDYHDNFTTSABEQQNREGDFYFMGFHAAGTAYYNSPSL 252
:
:
:
Qy 188 -----SKCYLTDLTINDSVSFFRTSKMYPHRPVLMVISHAAPGHEDS-AP-QYSR 237
:
:
:
Db 253 FKNERVPARGYISDQLTDEALGVVDRAKTL--DQPFMLYLAIVSAHLENDNPAEQYQK 310

[illegible]

QY 391 --DTERPVNREHL-----KKQKRVWRDSEFLVER--GK-----LLHKRDNDKVDQAE 432
Db 416 RPKGDKRTSOLYTPMPYGGQSYGQGVTRDYTLVIDRKVGKPLTYLTHDNKNDPYQMK 475
QY 433 ---ENFLPKYQVNDLQORAEYQYACQOLGQKQVCEATGKLUKLH 475
Db 476 IAAENMALVNQLIAD-----ELIPWLEHSGDVWRPTVEAANAAY 516

RESULT 15
KTHUAC
steryl-sulfatase (EC 3.1.6.2) precursor - human
N:Alternate names: arylsulfatase C; steroid sulfatase (STS); steryl-sulfate sulfohydrolase
C:Species: Homo sapiens (man)
C:Date: 21-May-1990 #sequence revision 27-Oct-1995 #text change 11-May-2000
C:Accession: A32641; A57116; A25961; S05415; S05423; I52800; I65619
J:Stein, C.; Hille, A.; Seidel, J.; Bijnhout, S.; Waheed, A.; Schmidt, B.; Geuze, H.; van
J. Biol. Chem. 264, 13865-13872, 1989
A:Title: Cloning and expression of human steroid-sulfatase. Membrane topology, glycosylation
A:Reference number: A32641; MUID:89340479; PMID:2668275
A:Accession: A32641
A:Molecule type: mRNA
A:Residues: 1-583 <STE>
A:Cross-references: GB:J04964; NID:G338564; PIDN:AAA60597.1; PID:G338565
A:Experimental source: BHK-21 cells
A:Note: Parts of this sequence were determined by protein sequencing
R:Yen, P.H.
unpublished results 1988, cited by GenBank
A:Reference number: A57116
A:Accession: A57116
A:Molecule type: mRNA
A:Residues: 1-22, 'E', 24-583 <YEN1>
A:Cross-references: GB:M16505; NID:G338513; PIDN:AAA60596.1; PID:G338514
R:Yen, P.H.; Allen, E.; Marsh, B.; Mohandas, T.; Wang, N.; Taggart, R.T.; Shapiro, L.J.
Cell 49, 443-454, 1987
A:Title: Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of
A:Reference number: A25961; MUID:87187642; PMID:3032454
A:Accession: A25961
A:Molecule type: mRNA
A:Residues: 1-22, 'E', 24-456, 'LRTTHPSGRFPSSPTSTFWFQRIALPHTCASVSGWSPITTHLYSLFF', 'KIPERE
A:Cross-references: GB:M16505; NID:G338513
A:Note: This sequence revised in A57116
R:Kawano, J.I.; Kotani, T.; Ohnaki, S.; Minamoto, N.; Matsuo, H.; Oinuma, T.; Aikawa, E.
Biochim. Biophys. Acta 997, 199-205, 1989
A:Title: Characterization of rat and human steroid sulfatases.
A:Reference number: S05414; MUID:89352671; PMID:2765556
A:Accession: S05415
A:Molecule type: protein
A:Residues: 22-43, 'X', 45-46 <KAW>
R:Dibbelt, L.; Otto, J.; Kuss, E.
Biol. Chem. Hoppe-Seyler 370, 847-848, 1989
A:Title: The N-terminal amino-acid sequence of human placental steryl-sulfatase.
A:Reference number: S05423; MUID:90074181; PMID:2590467
A:Accession: S05423
A:Molecule type: protein
A:Residues: 22-35 <DIB>
R:Yen, P.H.; Marsh, B.; Allen, E.; Teasi, S.P.; Ellison, J.; Connolly, L.; Neiswanger, K.
Cell 55, 1123-1135, 1989
A:Title: The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence
A:Reference number: I52800; MUID:89077541; PMID:3203382
A:Accession: I52800
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 134-274 <YEN3>
A:Cross-references: GB:M23945; NID:G338604; PIDN:AAA60598.1; PID:G338607
A:Accession: I65619
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 461-583 <YEN4>
A:Cross-references: GB:M23556; NID:G338605; PIDN:AAA60599.1; PID:G338608
C:Genetics:

A:Gene: GDB:STS
A:Cross-references: GDB:120393; OMIM:308100
A:Map position: Xp22.32-Xp22.32
A:Note: defects in this gene can cause X-linked ichthyosis
C:Function:
A:Description: hydrolyzes 3beta-hydroxysteroid sulfates to release sulfate
C:Superfamily: animal sulfatase
C:Keywords: endoplasmic reticulum; glycoprotein; lysosome; microsomal; sulfuric ester hyd
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-583/Product: steryl-sulfatase #status predicted <MAT>
F:185-211/Domain: transmembrane #status predicted <TM1>
F:213-237/Domain: transmembrane #status predicted <TM2>
F:47,259/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:75/Modified site: 3-oxoalanine (Cys) #status predicted
F:333,459/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.0%; Score 238.5; DB 1; Length 583;
Best Local Similarity 20.7%; Pred. No. 3.2e-09;
Matches 126; Conservative 67; Mismatches 180; Indels 237; Gaps 23;

QY 43 RPNILVLTDDQVELGSMQVM-NKTRRI-----MEQGAHFNAFVTPMCCPSRSIL 96
Db 26 RPNILVMAD--DLGIGDPCGCGNKTIPTNIDRLASGGVKLTQHLAASPLCTPSRAAFM 83
QY 97 TGYVHNHTYNNENCSSPSWQ-----AQHESRTPAVYLNSTGYTAFFGK 143
Db 84 TGRY-----PVRSGMASWSTGVFLFTASSGGLPTDEITFAKLLKQGGYSTALIGK 134
QY 144 Y-----LNEY-----NGSYVPPGWKE----- 159
Db 135 WHLGWSCHSKTDFCHPLHGFNYFYGISLTNLRDCKPGEVSFTTGFKLVFLQIVG 194
QY 160 -----WVGLLNSRFNYTLCRNGVKEHG 184
Db 195 VTLTLAALNCLGLHLVPLGVFFSLLFLAALILTLFLGFLHYPRPLNCFMWRN--YBIIQ 252
QY 185 SDYKDYLTDLITNDVSFFRTSKMYPHPVLMVISHAAPHGPEDSAPQYSLFFPNASQ 244
Db 253 QPMYDNLTLQBLTVEAAQFQRTNTE---TFLLVLVSLVHV-----TALPSSKDFAGKSO 304
QY 245 HTSPYNAFNPDKHIMRYTGMKPIHMEFTNNLQRLQTLMSVDDSMETIYNMLVET 304
Db 305 H--GVYGD-----VEEMDWSVQQLNLLDEL 329
QY 305 GELDNTYIVTADHGYHI-----GQFGLVK-GKSNPVEFDIRVPPYVGGPNV-E 351
Db 330 RLANDTLIYFTSDQGAHVEVSVSKGEIHGSGNGIYKGGKANNWEGGIRVPGILRPRVIQ 389
QY 352 AGCLNPHIVLNIDLAPTLIDAGLIDIPAD--MDGKSILKLLD--TERPVN----- 397
Db 390 AGQKIDEPTSNMIDIFTVAKLAGAPLPEDRIIDGRDLMPLEGGKSQSDHEFLPHYCNAY 449
QY 398 ---RPHLKKORVWRDSEFLVER-----GKLLHKRD-----NDKQDA 430
Db 450 LNAVVRHPQNSTSIWKAFFFTFNFNPVSGNCGFATHVCFGFSYVTHHDPFLFLDISKDP 509
QY 431 QEENFL-----PKTORVKDLQCRAE-----YQTACEQLGQK 461
Db 510 RERNPLTPASEPRPYEILKVMQEAADRHQTLPVPDPQFSMNNFLKWPWLQLCCPSTGLS 569
QY 462 WQVEDATGK 471
Db 570 CQCDREKQDK 579

Search completed: February 14, 2004, 22:58:59
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2004, 19:39:30 ; Search time 70 Seconds
(without alignments)
1972.745 Million cell used

Title: US-10-025-966A-6

Perfect score: 4750

Sequence: 1 MGPPSLVCLLSATVFSLIG.....PEMKRPSSKSLGQLWEGWEG 870

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Maximum DB seq length: 2000000000

[illegible]

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

```

Database :
A. Geneseq_19Jun03.*
1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4750	100.0	870	22	AAB85774	Human drug metabol
2	4750	100.0	870	22	AAW79215	Human protein SEQ
3	4750	100.0	870	22	AAE01440	Human gene 5' end
4	4750	100.0	870	23	ABG71836	human SUI72 protei
5	4750	100.0	870	23	ABG63903	Human albumin fusi
6	4746	99.9	870	22	AAE01471	Human gene 5' end
7	4746	99.9	870	23	ABE80921	Human sulphatase p
8	4746	99.9	870	23	ABG63902	Human albumin fusi
9	4746	99.9	870	23	ABB30378	Human polypeptide

10	4746	99.9	885	23	ABB80922	Human KIAA1427 pol
11	4723.5	99.4	867	21	AA56648	Membrane-bound pro
12	4723.5	99.4	867	22	AAU12181	Human PRO120 poly
13	4723.5	99.4	867	22	AAAB6171	Human PRO1120 (UNQ
14	4723.5	99.4	867	23	ABB95459	Human angiogenesis
15	4723.5	99.4	867	23	ABB84853	Human PRO1120 prot
16	4723.5	99.4	867	24	ABU68579	Human PRO polytype
17	4723.5	99.4	867	24	ABU68855	Human secreted/tra
18	4723.5	99.4	867	24	ABU53660	Novel secreted and
19	4723.5	99.4	867	24	ABU53064	Novel human secret
20	4723.5	99.4	867	24	ABU59211	Human secreted/tra
21	4723.5	99.4	867	24	ABU59360	Novel human secret
22	4723.5	99.4	867	24	ABU69495	Human secreted/tra
23	4723.5	99.4	867	24	ABU57986	Human PRO polytype
24	4723.5	99.4	867	24	ABU58917	Human secreted/tr
25	4723.5	99.4	867	24	ABU113877	Human PRO1120 poly
26	4723.5	99.4	867	24	ABU110832	Human PRO polytype
27	4616	97.2	850	22	AAW79216	Human protein SEQ
28	4473.5	94.2	875	23	ASG71838	Mouse SULF2 protei
29	4338	91.3	818	21	AAAB4268	Human OREF ORF032
30	3524	74.2	643	22	AAW80199	Human protein SEQ
31	3524	74.2	643	22	AAW80200	Human protein SEQ
32	3179.5	66.9	871	21	AAAB00191	Breast cancer prot
33	3179.5	66.9	871	22	AAAB85481	Human 23553 sulfat
34	3179.5	66.9	871	23	ASG71835	Human SULF1 protei
35	3179.5	66.9	871	23	ABG679640	Human secreted pro
36	3179.5	66.9	871	24	ABR47506	Breast cancer asso
37	3121.5	65.7	867	22	AAE00434	Quail sulfatase (Q
38	3063	64.5	818	23	AAU699417	Lung small cell ca
39	3063	64.5	1611	22	AAE00438	Human sulfatase (H
40	2977	62.7	800	22	AAAM25714	Human protein sequ
41	2692	56.8	490	22	AAW93919	Human polypeptide,
42	2692	56.7	490	22	AAW93864	Human polypeptide,
43	2356	49.6	434	22	AAE01522	Human gene 5' endc
44	2093	44.1	455	22	AAE00437	Mouse sulfatase (M
45	1718.5	36.2	1114	22	ABB71505	Drosophila melanop

ALIGNMENTS

RESULT 1	
AAB85774	
IID	AAB85774 standard; Protein; 870 AA.
XX	
XX	AAB85774;
XX	
XX	29-OCT-2001 (first entry)
XX	
DE	Human drug metabolizing enzyme (ID No. 1558210CD1).
XX	
KW	Drug metabolizing enzyme; DME; immunosuppressive; c
KW	hepatotropic; antiallergic; antiasthmatic; antibac
KW	antisense therapy; gene therapy; human.
XX	
OS	Homo sapiens.
XX	
PN	WO200159127-A2.
XX	
PD	16-AUG-2001.
XX	
XX	08-FEB-2001; 2001WO-US04423.
XX	
XX	11-FEB-2000; 2000US-0181856.
PPR	17-FEB-2000; 2000US-0183694.
XX	25-FEB-2000; 2000US-0185141.
PPR	03-MAR-2000; 2000US-0186818.
PPR	09-MAR-2000; 2000US-0188345.
PPR	17-MAR-2000; 2000US-0189997.
XX	
XX	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Tang YT. Yue H. Baughn MR. Yao MG. Bandman O.

PI Gandhi AR, Ring HZ, Shih LL, Yang J, Policky JL;
XX WPI; 2001-514673/56.
DR N-PSDB; AAH76201.
XX
XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the
PT diagnosis, treatment, and prevention of autoimmune/inflammatory, cell
PT proliferative, developmental and endocrine disorders -
XX
XX Claim 1; Page 133-134; 150pp; English.
XX
XX The invention provides human drug metabolizing enzymes (DME) and
CC polynucleotides encoding the DMEs. The DME can be expressed by standard
CC recombinant methodology. DMEs and their agonists and antagonists are
CC useful for the diagnosis, treatment, and prevention of autoimmune/
CC inflammatory, cell proliferative, developmental, endocrine such as
CC aneurysm, eye, metabolic, and gastrointestinal disorders, including liver
CC disorders and infection. The present sequence represents a human DME.
XX
SQ Sequence 870 AA;

Query Match 100.0%; Score 4750; DB 22; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVFLSGSSAPLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS 60
Db 1 MGPPSLVLCILSATVFLSGSSAPLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS 60

QY 61 MQVMNKRIRMEQGAHFINAFVITPMPCCPSRSILTKYVHNHTYTNNECSSPSWQA 120
Db 61 MQVMNKRIRMEQGAHFINAFVITPMPCCPSRSILTKYVHNHTYTNNECSSPSWQA 120

QY 121 QHESSTFAVLNSTGYRTAPFGKYLNEVNGSVPPGKWEVGLLNKSRFNYTLCRNGVK 180
Db 121 QHESSTFAVLNSTGYRTAPFGKYLNEVNGSVPPGKWEVGLLNKSRFNYTLCRNGVK 180

QY 181 EKHGSDYSKDYLTDLITNDSVFSFRTSKMYPHPRLVMI SHAAPHGPEDSAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVFSFRTSKMYPHPRLVMI SHAAPHGPEDSAPOYSRLFP 240

QY 241 NASQHTPSYNAFNPDKHMYRTGPKPIHMEFTNMQRLQTLMSVDDSDMETTYNM 300
Db 241 NASQHTPSYNAFNPDKHMYRTGPKPIHMEFTNMQRLQTLMSVDDSDMETTYNM 300

QY 301 LYVETGELDNITYIVYTADHGYHIGQGLVKGKSMPEYEDIRVPYVGPVNEAGCLNPHIV 360
Db 301 LYVETGELDNITYIVYTADHGYHIGQGLVKGKSMPEYEDIRVPYVGPVNEAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMDGKSLKLDTERPVRNFRHLKKOMRVNRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKSLKLDTERPVRNFRHLKKOMRVNRDSFLVERGKLL 420

QY 421 HKRDNDKVDQAEENFLPKYQVVKDLCQAEYQTAQCEQLGQKWCVEATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAEENFLPKYQVVKDLCQAEYQTAQCEQLGQKWCVEATGKLLHKCKGP 480

QY 481 MELGGSRLSNLVPKYVGGSEACTCDSDGYKL SLAGRRKLLFKKTKYKASVYRSRSIRSV 540
Db 481 MELGGSRLSNLVPKYVGGSEACTCDSDGYKL SLAGRRKLLFKKTKYKASVYRSRSIRSV 540

QY 541 ALEVDGRVYHVLGDAAPRNITKHWFGAPEDQDDKDGFGSTGGLPDYSAANPIKVT 600
Db 541 ALEVDGRVYHVLGDAAPRNITKHWFGAPEDQDDKDGFGSTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLDLKYSLQAWDKHLHDHETLQNKIKNLRVGRHGLKKRPEEC 660
Db 601 HRCYILENDTVQCDLDLKYSLQAWDKHLHDHETLQNKIKNLRVGRHGLKKRPEEC 660

QY 661 DCHKITSYHTQHGRLKHRGSSLHPFRKQLQEKDKVWLLREQKRRKLLKRLQNNNDTC 720
Db 661 DCHKITSYHTQHGRLKHRGSSLHPFRKQLQEKDKVWLLREQKRRKLLKRLQNNNDTC 720

721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNYWCMTINETHNLFCEPATGFE 780
721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNYWCMTINETHNLFCEPATGFE 780
781 YFDLNTDPQLMNAVNTLDRVNLQHLVQLMELRSCKYKQCNPRNMDLGLKDGGSYE 840
781 YFDLNTDPQLMNAVNTLDRVNLQHLVQLMELRSCKYKQCNPRNMDLGLKDGGSYE 840
841 QYRQFQRRKWPENKRPSSKSLGQLWEGWEG 870
841 QYRQFQRRKWPENKRPSSKSLGQLWEGWEG 870

RESULT 2
AAM79215
ID AAM79215 standard; Protein; 870 AA.
XX AAM79215;
XX AC
XX DT 06-NOV-2001 (first entry)
XX Human protein SEQ ID NO 1877.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.
XX Homo sapiens.
XX WO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US04098.
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX 15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
XX N-PSDB; AAK52348.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX Claim 20; Page 4262-4264; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM79215-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activity/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.

XX	DT	17-JUL-2001	(first entry)
XX	DE	Human gene 5 encoded secreted protein HBQN39, SEQ ID NO:95.	
XX	KW	Human; secreted protein; proliferative disorder; cancer; tumour;	
XX	KW	fetal abnormality; developmental abnormality; haematopoietic disorder;	
XX	KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;	
XX	KW	inflammation; allergy; neurological disorder; Alzheimer's disease;	
XX	KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;	
XX	KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;	
XX	KW	cardiovascular disorder; angiotensin II receptor; kidney disorder;	
XX	KW	gastrointestinal disorder; pregnancy-related disorder;	
XX	KW	endocrine disorder; infection; wound healing; vulnary;	
XX	KW	cell culture; chemotaxis; food additive; gene therapy;	
XX	OS	Homo sapiens.	
XX	PH	Key Location/Qualifiers	
XX	FT	Peptide 1..24	
XX	FT	Protein /label= Signal_peptide	
XX	FT	25..870	
XX	FT	/note= "Mature human secreted protein"	
XX	EN	WO200134626-A1.	
XX	PD	17-MAY-2001.	
XX	XX	01-NOV-2000; 2000WO-US30045.	
XX	PR	05-NOV-1999; 99US-0163581.	
XX	PR	30-JUN-2000; 2000US-0215133.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;	
XX	DR	N-PSDB; AAD05304.	
XX	PT	WPI; 2001-308778/32.	
XX	PT	New nucleic acid molecules encoding 28 human secreted proteins for	
XX	PT	diagnosing, preventing, treating or ameliorating medical conditions and	
XX	XX	used as food additives or preservatives -	
XX	PS	Claim 11; Page 486-489; 562pp; English.	
XX	CC	AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted	
XX	CC	protein genes, and AAE01436-AAE01513 represent the proteins they encode.	
XX	CC	The genes and their secreted proteins are useful for preventing,	
XX	CC	treatment or ameliorating medical conditions, e.g., by protein or gene	
XX	CC	therapy. Pathological conditions can be diagnosed by determining the	
XX	CC	amount of the new protein in a sample or by determining the presence of	
XX	CC	mutations in the new genes. Specific uses are described for each of the	
XX	CC	28 genes, based on the tissues in which they are most highly expressed,	
XX	CC	and include developing products for the diagnosis or treatment of	
XX	CC	proliferative disorders, cancer, tumors, foetal and developmental	
XX	CC	abnormalities, haematopoietic disorders, diseases of the immune system,	
XX	CC	AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,	
XX	CC	allergies, neurological disorders (e.g., Alzheimer's disease,	
XX	CC	Parkinson's disease), cognitive disorders, schizophrenia, asthma,	
XX	CC	skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,	
XX	CC	cardiovascular disorders, angiotensin II receptor, kidney disorders,	
XX	CC	gastrointestinal disorders, pregnancy-related disorders, endocrine	
XX	CC	disorders, and infections. The proteins can also be used to aid wound	
XX	CC	healing and epithelial cell proliferation, to prevent skin aging due to	
XX	CC	sunburn, to maintain organs before transplantation, for supporting cell	
XX	CC	culture of primary tissues, to regenerate tissues, to identify their	
XX	CC	cognate ligands or binding partners, and in chemotaxis, and can be used	
XX	CC	as a food additive or preservative to modify storage properties.	
XX	CC	Antibodies specific for a protein of the invention can be used in	
XX	CC	alleviating symptoms associated with the disorders mentioned above, and	

CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention.

XX	Sequence	870 AA;
SQ	Query Match	100.0%; Score 4750; DB 22; Length 870;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 870; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MGPPSLVCLLSATVSLGGSSAFSLHRLKGRFQDRNRIRNIIILVLTDDQVLEGS 60
DB	1	MGPPSLVCLLSATVSLGGSSAFSLHRLKGRFQDRNRIRNIIILVLTDDQVLEGS 60
QY	61	MQVMNKTIRIMEQGAHFINAFVTPMCCPSRSSILTKGYVHNHTYNNENCSPPSQWA 120
DB	61	MQVMNKTIRIMEQGAHFINAFVTPMCCPSRSSILTKGYVHNHTYNNENCSPPSQWA 120
QY	121	QHESTPAVILNSTGYRTAFPGKYLNEVNGSYVPGWKEWGLLKNSRFNYTLCRNGVK 180
DB	121	QHESTPAVILNSTGYRTAFPGKYLNEVNGSYVPGWKEWGLLKNSRFNYTLCRNGVK 180
QY	181	EKHGSDYKDYLTDLITDNDVSFFRTSKMYPHRPVLNVI SHAAPHGPDSDAPQYSLFP 240
DB	181	EKHGSDYKDYLTDLITDNDVSFFRTSKMYPHRPVLNVI SHAAPHGPDSDAPQYSLFP 240
QY	241	NASQHTPSYNAFNPDKHIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
DB	241	NASQHTPSYNAFNPDKHIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
QY	301	LVETGEIDNTYIVYADHGHIHQGLVKGKSMYPYEDIRVPFYVGNVVEAGCLNPHIV 360
DB	301	LVETGEIDNTYIVYADHGHIHQGLVKGKSMYPYEDIRVPFYVGNVVEAGCLNPHIV 360
QY	361	LNIDLAPTILDIAGLDIPADMGGKILKLLDTERPVNRFHLKQKQVVRDVSFLVERGKLL 420
DB	361	LNIDLAPTILDIAGLDIPADMGGKILKLLDTERPVNRFHLKQKQVVRDVSFLVERGKLL 420
QY	421	HKRDNDKVDQAQENFLPKYQVKDLCQAEVYQACEQLGQKQWQVEDATGKLKHKCKGP 480
DB	421	HKRDNDKVDQAQENFLPKYQVKDLCQAEVYQACEQLGQKQWQVEDATGKLKHKCKGP 480
QY	481	MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLPKKKYKASYVRSRSIRSV 540
DB	481	MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLPKKKYKASYVRSRSIRSV 540
QY	541	AIEVDGRVTHVGLGAAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB	541	AIEVDGRVTHVGLGAAQPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
QY	601	HRCYILENDTVQCDLLDYKSLQAWDKHLHIDHETIETLQNKIKNLRVGRHLKXKRPEEC 660
DB	601	HRCYILENDTVQCDLLDYKSLQAWDKHLHIDHETIETLQNKIKNLRVGRHLKXKRPEEC 660
QY	661	DCHKISYHTQKRLKRGSSLHPRFKGLQEKDKVWLLRQKQKXKRLKRLQNNDTIC 720
DB	661	DCHKISYHTQKRLKRGSSLHPRFKGLQEKDKVWLLRQKQKXKRLKRLQNNDTIC 720
QY	721	SMPLGTCFTHDQHWOTAPFWTLGPFCACTSANNNTYCMRTINETHFLCFEATGPLE 780
DB	721	SMPLGTCFTHDQHWOTAPFWTLGPFCACTSANNNTYCMRTINETHFLCFEATGPLE 780
QY	781	YFDLNTDTPQLMNAVNTLDRVLNQLHVQMLBRSCKGYKQCNPRTRNNDLGLKGGGYSY 840
DB	781	YFDLNTDTPQLMNAVNTLDRVLNQLHVQMLBRSCKGYKQCNPRTRNNDLGLKGGGYSY 840
QY	841	QYRQFORRWPEMKRPSKSLQLMWEGWG 870
DB	841	QYRQFORRWPEMKRPSKSLQLMWEGWG 870

ID	ABG71836 standard; Protein; 870 AA.
XX	ABG71836;
XX	20-JAN-2003 (first entry)
XX	human SULF2 protein.
DE	Human; SULF2; glucosamine-6-sulphatase; cancer; ischaemia; enzyme;
KW	tumour; angiogenesis; coronary; carotid; arterial occlusive disease;
KW	peripheral arterial disease; atherosclerosis; myointimal hyperplasia;
KW	thromboangitis obliterans; thrombotic disorder; vasculitis;
KW	heart attack; myocardial infarction; vascular death; inflammation;
KW	rheumatoid arthritis; asthma; adult respiratory distress syndrome;
KW	sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;
KW	allograft rejection; lymphoma; thrombosis; sulphatase.
OS	Homo sapiens.
XX	
XX	Key Location/Qualifiers
FT	Peptide 1..22 /note= "signal peptide"
FT	Protein 25..870 /note= "Mature protein"
FT	Cleavage-site 409..410 /note= "furan/PACE protease cleavage site"
FT	Cleavage-site 423..424 /note= "furan/PACE protease cleavage site"
FT	Cleavage-site 538..539 /note= "furan/PACE protease cleavage site"
FT	Cleavage-site 565..566 /note= "furan/PACE protease cleavage site"
XX	
XX	WO200259327-A2.
XX	01-AUG-2002.
XX	26-DEC-2001; 2001WO-US49793.
XX	27-DEC-2000; 2000US-258577P.
PR	09-FEB-2001; 2001US-267831P.
XX	(REGC) UNIV CALIFORNIA.
XX	Rosen SD, Hemmerich S, Tomita M, Palmeri D;
XX	WPI; 2002-636541/68.
DR	N-PSDB; ABS56553.
XX	New sulfatase polypeptides having glucosamine-6-sulfatase activity,
PT	useful in screening, discovery and preparation of diagnostic and
PT	therapeutic agents for treating cancer, ischemic conditions,
PT	inflammation, or thrombosis -
XX	Example 1; Fig2A; 293pp; English.
PS	This invention relates to the DNA and protein sequences of a novel
XX	polypeptide having glucosamine-6-sulphatase activity. The sulphatases of
CC	are useful in screening, discovery and preparation of diagnostic and
CC	therapeutic agents for treating cancer, ischemic conditions,
CC	inflammation, or thrombosis. The nucleic acids are useful in preparing
CC	the sulfatase polypeptides, identifying the expression of genes in a
CC	biological specimen, or generating transgenic non-human animals or
CC	site-specific gene modification in cell lines. The host cells are
CC	useful in replicating and/or expressing the polynucleotides or nucleic
CC	acids. The agents are useful in treating the disorders cited above by
CC	reducing tumour growth, inflammation, and thrombosis, or increasing
CC	angiogenesis, e.g. by treating coronary, carotid, or arterial occlusive
CC	disease, peripheral arterial disease, atherosclerosis, myointimal
CC	hyperplasia, thromboangitis obliterans, thrombotic disorders,
CC	vasculitis; or preventing ischemic conditions, heart attack (myocardial
CC	infarction), or other vascular death. The sulphatases and/or agents are
CC	also useful in treating rheumatoid arthritis, asthma, adult respiratory

CC	distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple
CC	sclerosis, allograft rejection, and spread of lymphomas to cutaneous
CC	sites. The present sequence represents the human SULF2 protein of the
CC	invention.
XX	
XX	
SQ	Sequence 870 AA;
Query Match	100.0%; Score 4750; DB 23; Length 870;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 870; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGFQDRNRIRNIIILVLTDDQDELGS 60
DB	1 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGFQDRNRIRNIIILVLTDDQDELGS 60
QY	61 MQVMNKTTRIMEOGGAHFINAFVITPMCCPSRSSILTKGVVHNHTYNNENCSFSSWQA 120
DB	61 MQVMNKTTRIMEOGGAHFINAFVITPMCCPSRSSILTKGVVHNHTYNNENCSFSSWQA 120
QY	121 QHESRTFAVILNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRPFNYTLCRNGVK 180
DB	121 QHESRTFAVILNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRPFNYTLCRNGVK 180
QY	181 EKHGSDYSKDYLTDLITNDSVSFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLFP 240
DB	181 EKHGSDYSKDYLTDLITNDSVSFRTSKMYPHRPVLMVISHAAPHGPEDSAPOYSRLFP 240
QY	241 NASQHITPSYNAIPDKHIMRTGPMKPIHMEFTNMLQKQLQTLMSVDDSMETIYNN 300
DB	241 NASQHITPSYNAIPDKHIMRTGPMKPIHMEFTNMLQKQLQTLMSVDDSMETIYNN 300
QY	301 LVEIGELDNTVIIVTADGHYHIGFGLVKGSMPEYEDIRVPFVVRGPNVEAGCLNPHIV 360
DB	301 LVEIGELDNTVIIVTADGHYHIGFGLVKGSMPEYEDIRVPFVVRGPNVEAGCLNPHIV 360
QY	361 LNI DLAPTILDIAGLDPADMDGKSIILKLDTERPVNRFLKQORVWRDSFLVERGKLL 420
DB	361 LNI DLAPTILDIAGLDPADMDGKSIILKLDTERPVNRFLKQORVWRDSFLVERGKLL 420
QY	421 HRRNDKVDQAEENFLPKYQVVKDLCORAEVQTACEQLGQKQWQVEDATGKLLHKCKGP 480
DB	421 HRRNDKVDQAEENFLPKYQVVKDLCORAEVQTACEQLGQKQWQVEDATGKLLHKCKGP 480
QY	481 MRLGGSRALSNLVPKYQGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
DB	481 MRLGGSRALSNLVPKYQGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
QY	541 AIEVDGRVYHVLGDAQPRNLTKRHWFGAPEDQDDKGGDFSGTGLPDYSAAPIKVT 600
DB	541 AIEVDGRVYHVLGDAQPRNLTKRHWFGAPEDQDDKGGDFSGTGLPDYSAAPIKVT 600
QY	601 HRCVILENDTVQCDLDLYKSLQAKDKHLHDHIEITLQNKILNLRVRGHLKKRPEEC 660
DB	601 HRCVILENDTVQCDLDLYKSLQAKDKHLHDHIEITLQNKILNLRVRGHLKKRPEEC 660
QY	661 DCHKISYHTQHGRLKRGSSLSLHFRKGLQEKDKVLLRLEQKQKLLKRLQNNQDTC 720
DB	661 DCHKISYHTQHGRLKRGSSLSLHFRKGLQEKDKVLLRLEQKQKLLKRLQNNQDTC 720
QY	721 SMPGLTCTHNDONHQTAPFWTLGPFCACTSANNTYWCMTNETHNLFCFATGFLF 780
DB	721 SMPGLTCTHNDONHQTAPFWTLGPFCACTSANNTYWCMTNETHNLFCFATGFLF 780
QY	781 YFDLNTDPYQLMNAVNTLDRVLNQLHVLQMLRSCSKGYKQCPRTNMDLGLKDGGSYE 840
DB	781 YFDLNTDPYQLMNAVNTLDRVLNQLHVLQMLRSCSKGYKQCPRTNMDLGLKDGGSYE 840
QY	841 QYRQFORRWKPEMKRPSKSLGQLWEGWEG 870
DB	841 QYRQFORRWKPEMKRPSKSLGQLWEGWEG 870

RESULT 5

ABG63903
ID ABG63903 standard; Protein; 870 AA.

XX
AC ABG63903;

XX
DT 27-AUG-2002 (first entry)

XX
DE Human albumin fusion protein #578.

XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antitumor;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurologic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.

XX
OS Homo sapiens.

OS
Synthetic.

XX
PN WO200177137-A1.

XX
PD 18-OCT-2001.

XX
PF 12-APR-2001; 2001WO-US11988.

XX
PR 12-APR-2000; 2000US-229358P.

PR
25-APR-2000; 2000US-199384P.

PR
21-DEC-2000; 2000US-256931P.

XX
(HUMA-) HUMAN GENOME SCI INC.

XX
PI Rosen CA, Haseltine WA;

XX
WPI; 2002-010886/01.

XX
PT New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -

XX
PS Claim 1; Page 892-894; 2102pp; English.

XX
CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC fusion proteins of the invention.

XX
SQ Sequence 870 AA;

Query Match 100.0%; Score 4750; DB 23; Length 870;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGFQDRNRIRNIIILVLTDDQDELGS 60

DB 1 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGFQDRNRIRNIIILVLTDDQDELGS 60

QY 61 MQVMNKTTRIMEOGGAHFINAFVITPMCCPSRSSILTKGVVHNHTYNNENCSFSSWQA 120

DB 61 MQVMNKTTRIMEOGGAHFINAFVITPMCCPSRSSILTKGVVHNHTYNNENCSFSSWQA 120

QY 121 QHESRTFAVILNSTGYRTAFPGKYLNEYSGYVPPGKWEVGLLKNRPFNYTLCRNGVK 180

Db 121 QHESRTFAVLNLTGRTAFTGKYLNEYNGSVPPGKWEVGLLKNSGRFYNYTLCRNGVK 180
Qy 181 EKHGSDYKDYLTDLITNDYSVFRTSKMYPHRPVLMVISHAAPHGPEDSAPQYSLRFP 240
Db 181 EKHGSDYKDYLTDLITNDYSVFRTSKMYPHRPVLMVISHAAPHGPEDSAPQYSLRFP 240
Qy 241 NASQHITTSYNAFNPDKHWMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASQHITTSYNAFNPDKHWMYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELNTYIVYTADHGHIHQGLVKGSMPEYEDIRVPFVYRGPNVEAGCLNHIV 360
Db 301 LVETGELNTYIVYTADHGHIHQGLVKGSMPEYEDIRVPFVYRGPNVEAGCLNHIV 360
Qy 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXKVRVWRDSEFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXKVRVWRDSEFLVERGKLL 420
Qy 421 HKRDNDKVDQAEENFLPKYQVVDLCQAEYQVQACEQLGQKQCVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAEENFLPKYQVVDLCQAEYQVQACEQLGQKQCVEDATGKLLHKCKGP 480
Qy 481 MRLGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKXLFKKYKASVVRGSRISV 540
Db 481 MRLGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKXLFKKYKASVVRGSRISV 540
Qy 541 AIEVDGRVYHVLGDAAPRNLTKEHWPAPEDQDDXGDFSGTGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAAPRNLTKEHWPAPEDQDDXGDFSGTGLPDYSAANPIKVT 600
Qy 601 HRCYILENDTVQCDLDLYKSLQAKWDHKLHIDHETETIQLNKILREVRGHLKKRPEEC 660
Db 601 HRCYILENDTVQCDLDLYKSLQAKWDHKLHIDHETETIQLNKILREVRGHLKKRPEEC 660
Qy 661 DCHKLSYTHQHGRLKHGSSLHPRFKGLQEKDKVYMLREQKXKGLKRLQNNNDTC 720
Db 661 DCHKLSYTHQHGRLKHGSSLHPRFKGLQEKDKVYMLREQKXKGLKRLQNNNDTC 720
Qy 721 SMPGLTCFTHDQHQWQTAFTWTLGPFCACTSANNTYTCMTINETHNPLFCBATGPLE 780
Db 721 SMPGLTCFTHDQHQWQTAFTWTLGPFCACTSANNTYTCMTINETHNPLFCBATGPLE 780
Qy 781 YFDLNTDTPQLMNAVNTLDRVLNLHVLQMLRSCGYKQCNPTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDTPQLMNAVNTLDRVLNLHVLQMLRSCGYKQCNPTRNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLQGLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLQGLWEGWEG 870

RESULT 6
AAE01471
ID AAE01471 standard; Protein; 870 AA.
XX AAE01471;
AC AAE01471;
XX AAE01471;
DT 17-JUL-2001 (first entry)
XX Human gene 5 encoded secreted protein HBQ39, SEQ ID NO:127.
DE Human;
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angioinfection; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulnerability;
XX cell culture; chemotaxis; food additive; gene therapy;
XX binding partner identification.

OS Homo sapiens.
XX Key Location/Qualifiers
XX Peptide 1..24
XX FT /label= Signal_peptide
XX FT 25..870
XX FT /note= "Mature human secreted protein"
XX FT Misc-difference 537
XX FT /label= Unknown
XX FT /note= "Encoded by ART"
XX WC200134626-A1.
XX 17-MAY-2001.
XX 01-NOV-2000; 2000WC-US30045.
XX 05-NOV-1999; 99US-0163581.
XX 30-JUN-2000; 2000US-0215133.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;
XX WPI; 2001-308778/32.
XX N-PSDB; AAD01471.
XX New nucleic acid molecules encoding 28 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives -
XX Claim 11; Page 507-510; 562pp; English.
XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
XX protein genes, and AAE01436-AAE01513 represent the proteins they encode.
XX AAE01514-AAE01544 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 28 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angioinfection, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein of the invention.
XX SQ Sequence 870 AA;
Query Match 99.9%; Score 4746; DB 22; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGPPSLVCLLSATVPSLLGGSAFLSHRLKGRFORDRNRINPNILVTTDDQVELGS 60
Db 1 MGPPSLVCLLSATVPSLLGGSAFLSHRLKGRFORDRNRINPNILVTTDDQVELGS 60

```
QY 61 MQVMKTRIMEQGAHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNENCSSPSWQA 120
Db 61 MQVMKTRIMEQGAHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNENCSSPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGKWEVGLLKNRFFNYTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDVSVFRTSKMYPHRPVLVISHAAPHGPEPSAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDVSVFRTSKMYPHRPVLVISHAAPHGPEPSAPOYSRLFP 240
QY 241 NASQHTPSYNAPNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASQHTPSYNAPNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
QY 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
Db 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKLLFKKYYKASVYRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYQGGSEACTCDSDGYKLSLAGRRKLLFKKYYKASVYRSRSIRSV 540
QY 541 ALEVDRVYHVLGDAAPRNLTKRHWFGAPEDODDQKGGDFSGTGGLPDYSAANPKVT 600
Db 541 ALEVDRVYHVLGDAAPRNLTKRHWFGAPEDODDQKGGDFSGTGGLPDYSAANPKVT 600
QY 601 HRCYILENDTVQCDLIDLYKSLQAWKDKHLHDHIEITLQNKIKNLRVGRHGLKKRPEEC 660
Db 601 HRCYILENDTVQCDLIDLYKSLQAWKDKHLHDHIEITLQNKIKNLRVGRHGLKKRPEEC 660
QY 661 DCHKLSYTHQKGRUKHRSLSLHPRKGLQKDKWLLREQRKKLRLKLLKRLQNNDC 720
Db 661 DCHKLSYTHQKGRUKHRSLSLHPRKGLQKDKWLLREQRKKLRLKLLKRLQNNDC 720
QY 721 SMPGLTCTFDHNOHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780
Db 721 SMPGLTCTFDHNOHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780
QY 781 YPDLNTDYPQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKOCNPRTRNMOLGLKDGGSYE 840
Db 781 YPDLNTDYPQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKOCNPRTRNMOLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPPSSKSLQOLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPPSSKSLQOLWEGWEG 870
RESULT 7
ABB80921
ID ABB80921 standard; Protein; 870 AA.
XX AC ABB80921;
XX DT 08-OCT-2002 (first entry)
XX DE Human sulphatase polypeptide, 22437.
XX KW Human; sulphatase; 22437; cytostatic; vulnerary; neuroprotective;
XX OS Gene therapy; enzyme.
XX OS Homo sapiens.
XX PN WO200252019-A2.
```

```
XX 04-JUL-2002.
XX 03-OCT-2001; 2001WO-US30856.
XX 21-DEC-2000; 2000US-257082P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Glucksmann MA, Rudolph-Owen LA;
XX WPI; 2002-566677/60.
XX N-PSDB; ABB86352, ABB86353.
XX Identifying agents for modulating (e.g. inhibiting) e.g. tumor
XX establishment, growth or metastases, neuron growth, or wound healing by
XX determining whether a test compound binds with a 22437 polypeptide
XX (human sulfatase) -
XX Claim 19; Fig 1A-F; 143pp; English.
XX The invention relates to identifying a compound useful for modulating at
XX least one phenomenon (e.g. tumour establishment, tumour growth, tumour
XX metastases, epithelial and/or endothelial cell proliferation, neuronal
XX cell growth, wound healing or cerebral injury). The method involves
XX determining whether a test compound binds with a 22437 polypeptide.
XX The identified modulators of 22437 nucleic acid and polypeptide are also
XX useful for treating cancer or wounds (e.g. stroke-related cerebral
XX ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid
XX and polypeptide are useful for diagnosing, preventing or treating a
XX subject having cancer or a cellular proliferation and/or differentiation
XX disorder or at risk of developing cancer or a cellular proliferation
XX and/or differentiation disorder. The present sequence represents the
XX human sulphatase polypeptide, 22437.
XX Sequence 870 AA;
Query Match 99.9%; Score 4746; DB 23; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGPPSLVLCCLLSATVFSLLGGSSAFSLSHRLKGRFQDRRIRNPNIILVTTDDQDVELGS 60
Db 1 MGPPSLVLCCLLSATVFSLLGGSSAFSLSHRLKGRFQDRRIRNPNIILVTTDDQDVELGS 60
QY 61 MQVMKTRIMEQGAHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNENCSSPSWQA 120
Db 61 MQVMKTRIMEQGAHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNENCSSPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGKWEVGLLKNRFFNYTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFGKYLNEVNGSVPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDVSVFRTSKMYPHRPVLVISHAAPHGPEPSAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDVSVFRTSKMYPHRPVLVISHAAPHGPEPSAPOYSRLFP 240
QY 241 NASQHTPSYNAPNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASQHTPSYNAPNPDKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGKSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKSIKLLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
QY 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
Db 421 HKRDNKVDQAQENFLPKYQVRKDLQRAEYQTACEQLGQKWCQVEDATGKLLHKCKGP 480
```

QY 481 MRLGSRALSNLVPRYVYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540
DB 481 MRLGSRALSNLVPRYVYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540
QY 541 ALEVDGRVYHVLGDAAPRNLTQRHWPCAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
DB 541 ALEVDGRVYHVLGDAAPRNLTQRHWPCAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
QY 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHETETLQNKIKNLRVGRHLKKRPEEC 660
DB 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHETETLQNKIKNLRVGRHLKKRPEEC 660
QY 661 DCHKISYHTQHKGRKLRHGRSSLHPPFRKGLQEKDKVWLLREQRKKLRLKRLQNNDT 720
DB 661 DCHKISYHTQHKGRKLRHGRSSLHPPFRKGLQEKDKVWLLREQRKKLRLKRLQNNDT 720
QY 721 SMPLGTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCBPATGFL 780
DB 721 SMPLGTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCBPATGFL 780
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNMDLGLKDGGSYE 840
QY 841 QYRQFORRWPEMKRPPSSKSLQLWEGWEG 870
DB 841 QYRQFORRWPEMKRPPSSKSLQLWEGWEG 870

RESULT 8

ID ABG63902
XX ABG63902 standard; Protein; 870 AA.
AC ABG63902;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #577.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; anti-infectivity; anti-inflammatory; antitumor;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US11988.
XX
PR 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
PI WPI; 2002-010886/01.
XX
DR
XX
PT New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -
XX
PS Claim 1; Page 889-891; 2102pp; English.
XX
CC The present invention relates to albumin fusion proteins comprising a

therapeutic protein X and human albumin (HA, also known as human serum albumin, HSA). The proteins are useful for treating a disease or disorder that may be modulated by therapeutic protein X. The albumin extends the shelf-life of protein X, and may increase its biological in vitro/in vivo activity. The protein is useful for treating and diagnosing disorders such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes), haematopoietic disorders, neural disorders (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis, schizophrenia), and connective disorders (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin fusion proteins of the invention.

XX Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 23; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLLSATVPSLLGSSAFSLSHRLKGRFQDRRNIRPNILVLTDDQDVELGS 60
DB 1 MGPPSLVLCCLLSATVPSLLGSSAFSLSHRLKGRFQDRRNIRPNILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEQGGAHFINAVFTTMCPCSSSILTKGYVHNHTYTNNECSSPSQA 120
DB 61 MQVMKTRIMEQGGAHFINAVFTTMCPCSSSILTKGYVHNHTYTNNECSSPSQA 120
QY 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYNGSVVPPGKWEVGLLKNSRPFYNTLCRNGYK 180
DB 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYNGSVVPPGKWEVGLLKNSRPFYNTLCRNGYK 180
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQYPHRVLAVISHAAPHGSDSPQSRLEPP 240
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKQYPHRVLAVISHAAPHGSDSPQSRLEPP 240
QY 241 NASOHTPSYNYAPNPDKHWIMRYTGPMPKPIHMETNMLQRKRLQTLMSVDDSMETIYNN 300
DB 241 NASOHTPSYNYAPNPDKHWIMRYTGPMPKPIHMETNMLQRKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGBELNTYIVYADHGYHIGQFGLVKGKSNPVEFDIRVPFYVRGPNVEAGCLNPHIV 360
DB 301 LVETGBELNTYIVYADHGYHIGQFGLVKGKSNPVEFDIRVPFYVRGPNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSLILKLLDTERPVRPHLKKQKVRWDSFLVERGKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSLILKLLDTERPVRPHLKKQKVRWDSFLVERGKLL 420
QY 421 HKRDNDKVDABEENFLPKYQVRVKDLQRAEYQTAECQLGQKQCCVEDATGKLLHKCKGP 480
DB 421 HKRDNDKVDABEENFLPKYQVRVKDLQRAEYQTAECQLGQKQCCVEDATGKLLHKCKGP 480
QY 481 MRLGSRALSNLVPRYVYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540
DB 481 MRLGSRALSNLVPRYVYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV 540
QY 541 ALEVDGRVYHVLGDAAPRNLTQRHWPCAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
DB 541 ALEVDGRVYHVLGDAAPRNLTQRHWPCAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
QY 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHETETLQNKIKNLRVGRHLKKRPEEC 660
DB 601 HRCYILENDTVQCDLDDLYKSLQAWKDKHLHDHETETLQNKIKNLRVGRHLKKRPEEC 660
QY 661 DCHKISYHTQHKGRKLRHGRSSLHPPFRKGLQEKDKVWLLREQRKKLRLKRLQNNDT 720
DB 661 DCHKISYHTQHKGRKLRHGRSSLHPPFRKGLQEKDKVWLLREQRKKLRLKRLQNNDT 720
QY 721 SMPLGTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCBPATGFL 780
DB 721 SMPLGTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCBPATGFL 780
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNMDLGLKDGGSYE 840

Db 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKGGGSYE 840
QY 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870
Db 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870

RESULT 9
ABB90378
ID ABB90378 standard; Protein; 870 AA.
XX
AC ABB90378;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2754.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW aniallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
XX
DE N-PSDB; ABL90787.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -
Claim 11; SEQ ID NO 2754; 2081pp + Sequence Listing; English.

The invention relates to novel genes (ABL90449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 870 AA;
Query Match 99.9%; Score 4746; DB 23; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLLSATVFSLLGGSSAFSLSHRLKGRFORDRNRIRPNIIILVLTDDQDVELGS 60
Db 1 MGPPSLVLCCLLSATVFSLLGGSSAFSLSHRLKGRFORDRNRIRPNIIILVLTDDQDVELGS 60
QY 61 MOVNKTTRIMEQGAHFINAFVTTMCCPSSSSILTQKYVHNHTYNNENCSPPSQA 120
Db 61 MOVNKTTRIMEQGAHFINAFVTTMCCPSSSSILTQKYVHNHTYNNENCSPPSQA 120
QY 121 QHESRTFAVYLNSTGYRTAFGKYLYNEVNGSVPPGKWEVGLLKNRSFYNTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFGKYLYNEVNGSVPPGKWEVGLLKNRSFYNTLCRNGVK 180
QY 181 EKXGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLVPMVISHAAPHGPEDESAPOYSRLFF 240
Db 181 EKXGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLVPMVISHAAPHGPEDESAPOYSRLFF 240
QY 241 NASOHITPSYVAPNPDKHMRYTGMKPIHMETNMLORKLOTLMSVDDSMETIYNN 300
Db 241 NASOHITPSYVAPNPDKHMRYTGMKPIHMETNMLORKLOTLMSVDDSMETIYNN 300
QY 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSMPEYFDIRVPFVYRGPNVEAGCLNPHIV 360
Db 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSMPEYFDIRVPFVYRGPNVEAGCLNPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKILKLDTERPVNRFHLKKQKRVWRDPLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKILKLDTERPVNRFHLKKQKRVWRDPLVERGKLL 420
QY 421 HKRDNKVDAAQENFLPKYQKVKDLCCRAEYQTAQCEQLGQKQWQCVEDATGKLKHKCKGP 480
Db 421 HKRDNKVDAAQENFLPKYQKVKDLCCRAEYQTAQCEQLGQKQWQCVEDATGKLKHKCKGP 480
QY 481 MRLGGSRALSNLVPKYQGSEACTCDSGYKLSLAGRRKKLKKKYYKASYVRXRSIRSV 540
Db 481 MRLGGSRALSNLVPKYQGSEACTCDSGYKLSLAGRRKKLKKKYYKASYVRXRSIRSV 540
QY 541 ALEVDGRVYHVLGDAQAPRNLTKEHWPAGPEDDDDXGDFSGTGGLPDYSAANPIKVT 600
Db 541 ALEVDGRVYHVLGDAQAPRNLTKEHWPAGPEDDDDXGDFSGTGGLPDYSAANPIKVT 600
QY 601 HRCYLENDTVQCOLDLYKSLQAWKHKLHIDHIEITLQNKIKNLRVGRHLKKRPREEC 660
Db 601 HRCYLENDTVQCOLDLYKSLQAWKHKLHIDHIEITLQNKIKNLRVGRHLKKRPREEC 660
QY 661 DCHKISYTHQKRLKRGSSLHPFRKGLQKDKVLLREOKRKKKLLKRLQNNDDTC 720
Db 661 DCHKISYTHQKRLKRGSSLHPFRKGLQKDKVLLREOKRKKKLLKRLQNNDDTC 720
QY 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFL 780
Db 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFL 780
QY 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKGGGSYE 840
Db 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKGGGSYE 840
QY 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870
Db 841 QYRQFQRKWPCKRPPSSKSLGQLWEGWEG 870
RESULT 10
ABB80922
ID ABB80922 standard; Protein; 885 AA.
XX
AC ABB80922;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human KIAA1427 polypeptide.
XX
KW Human; sulphatase; 22437; cytostatic; vulnary; neuroprotective;
KW gene therapy; KIAA1427.

OS Homo sapiens.
XX WO2000252019-A2.
XX
XX
XX PD 04-JUL-2002.
XX
XX PF 03-OCT-2001; 2001WO-US30856.
XX
XX PR 21-DEC-2000; 2000US-257082P.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX FI Glucksmann MA, Rudolph-Owen LA;
XX
XX DR WPI; 2002-566677/60.
XX
XX DR N-PSDB; AEN86354.
XX
XX PT Identifying agents for modulating (e.g. inhibiting) e.g. tumor
XX establishment, growth or metastases, neuron growth, or wound healing by
XX determining whether a test compound binds with a 22437 polypeptide
XX (human sulfatase) -
XX
XX PS Disclosure; Fig 4A-D; 143pp; English.
XX
XX CC The invention relates to identifying a compound useful for modulating at
XX least one phenomenon (e.g. tumour establishment, tumour growth, tumour
XX metastases, epithelial and/or endothelial cell proliferation, neuronal
XX cell growth, wound healing or cerebral injury). The method involves
XX determining whether a test compound binds with a 22437 polypeptide.
XX The identified modulators of 22437 nucleic acid and polypeptide are also
XX useful for treating cancer or wounds (e.g. stroke-related cerebral
XX ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid
XX and polypeptide are useful for diagnosing, preventing or treating a
XX subject having cancer or a cellular proliferation and/or differentiation
XX disorder or at risk of developing cancer or a cellular proliferation
XX and/or differentiation disorder. The present sequence represents a human
XX KIAA1427 polypeptide used in alignment studies with the human sulphatase
XX polypeptide, 22437.
XX
XX SQ Sequence 885 AA;

Query Match 99.9%; Score 4746; DB 23; Length 885;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNIILVLTDDQVELGS 60
DB 16 MGPPSLVLCILSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNIILVLTDDQVELGS 75
QY 61 MQVMNKRIRIMEQGAHFINAFVTPMCCPSRSSILTKGVVHNHTYNNENCSPPSWQA 120
DB 76 MQVMNKRIRIMEQGTHTFINAFVTPMCCPSRSSILTKGVVHNHTYNNENCSPPSWQA 135
QY 121 QHESRTFAYLNSGYRTAFPGKYLINEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
DB 136 QHESRTFAYLNSGYRTAFPGKYLINEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 195
QY 181 EKHGSDYSKDLVLTDLTNDVSVPFRTSKMYPHPVPLVWLSHAHPGPDSPAQYSLRFP 240
DB 196 EKHGSDYSKDLVLTDLTNDVSVPFRTSKMYPHPVPLVWLSHAHPGPDSPAQYSLRFP 255
QY 241 NASQHITPSYNAFPDCKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 256 NASQHITPSYNAFPDCKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 315
QY 301 LVEGTCLDNTVIVYADGHYHIGFGLVKGSMEYEPDIRVPFYVRGPNVEAGCLAPHIV 360
DB 316 LVEGTCLDNTVIVYADGHYHIGFGLVKGSMEYEPDIRVPFYVRGPNVEAGCLAPHIV 375
QY 361 LNIDLAPITLDIADIDIPADMGGKSLKLDTERPVNRFHLKKQVRWRDSFLVERGKLL 420
DB 376 LNIDLAPITLDIADIDIPADMGGKSLKLDTERPVNRFHLKKQVRWRDSFLVERGKLL 435

QY 421 HKRDNDKYDAQENFLPKYQVRKDLCPRAEYQTACEQLGQKWCQVEDATGKLKHKCKGP 480
DB 436 HKRDNDKYDAQENFLPKYQVRKDLCPRAEYQTACEQLGQKWCQVEDATGKLKHKCKGP 495
QY 481 MRLGGSRALSNLVPKYQGSGSEACTCDSDGYKLSLAGRRKLFKKYKASVYRSRSRSV 540
DB 496 MRLGGSRALSNLVPKYQGSGSEACTCDSDGYKLSLAGRRKLFKKYKASVYRSRSRSV 555
QY 541 ALEVDGRVYHVLGDAAOPRNLTKRHWPAGAPEDODKGGDFSGTGGLPDYSAANPIKVT 600
DB 556 ALEVDGRVYHVLGDAAOPRNLTKRHWPAGAPEDODKGGDFSGTGGLPDYSAANPIKVT 615
QY 601 HRCYILENDTVQCCLDLYKSLQAWKHKLHIDHIEITLQNKIKNLREYRHLKKRPEEC 660
DB 616 HRCYILENDTVQCCLDLYKSLQAWKHKLHIDHIEITLQNKIKNLREYRHLKKRPEEC 675
QY 661 DCHKISYHTQHKGLKHRGSSLHPFRKGLQEKVKVLLREQKRLKRLKRLQNDTC 720
DB 676 DCHKISYHTQHKGLKHRGSSLHPFRKGLQEKVKVLLREQKRLKRLKRLQNDTC 735
QY 721 SMPGLTCETHDNQHWQAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780
DB 736 SMPGLTCETHDNQHWQAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 795
QY 781 YFDLNTDTPYQLMNAVNTLDRVLNQLHVQLMELRSCKGYKOCNPRTRNMDLGLKDGGSYE 840
DB 796 YFDLNTDTPYQLMNAVNTLDRVLNQLHVQLMELRSCKGYKOCNPRTRNMDLGLKDGGSYE 855
QY 841 QYRQFORAKWPEMKRPPSSKSLGQWEGWEG 870
DB 856 QYRQFORAKWPEMKRPPSSKSLGQWEGWEG 885

RESULT 11
AAY66648
ID AAY66648 standard; protein; 867 AA.
XX
AC AAY66648;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1120.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping.
XX
OS Homo sapiens.
XX
PN WC9963088-A2.
XX
PD
XX
PF 09-DEC-1999.
XX
FF 02-JUN-1999; 99WO-US12252.
XX
XX 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.

PR 10-JUN-1998; 98US-0088734.
 PR 10-JUN-1998; 98US-0088738.
 PR 10-JUN-1998; 98US-0088740.
 PR 10-JUN-1998; 98US-0088741.
 PR 10-JUN-1998; 98US-0088742.
 PR 10-JUN-1998; 98US-0088810.
 PR 10-JUN-1998; 98US-0088811.
 PR 10-JUN-1998; 98US-0088824.
 PR 10-JUN-1998; 98US-0088825.
 PR 10-JUN-1998; 98US-0088826.
 PR 11-JUN-1998; 98US-0088858.
 PR 11-JUN-1998; 98US-0088861.
 PR 11-JUN-1998; 98US-0088863.
 PR 11-JUN-1998; 98US-0088876.
 PR 12-JUN-1998; 98US-0089090.
 PR 12-JUN-1998; 98US-0089105.
 PR 16-JUN-1998; 98US-0089440.
 PR 16-JUN-1998; 98US-0089512.
 PR 16-JUN-1998; 98US-0089514.
 PR 17-JUN-1998; 98US-0089532.
 PR 17-JUN-1998; 98US-0089538.
 PR 17-JUN-1998; 98US-0089598.
 PR 17-JUN-1998; 98US-0089599.
 PR 17-JUN-1998; 98US-0089600.
 PR 17-JUN-1998; 98US-0089653.
 PR 18-JUN-1998; 98US-0089801.
 PR 18-JUN-1998; 98US-0089907.
 PR 18-JUN-1998; 98US-0089908.
 PR 19-JUN-1998; 98US-0089947.
 PR 19-JUN-1998; 98US-0089948.
 PR 19-JUN-1998; 98US-0089952.
 PR 22-JUN-1998; 98US-0090246.
 PR 22-JUN-1998; 98US-0090252.
 PR 23-JUN-1998; 98US-0090254.
 PR 23-JUN-1998; 98US-0090349.
 PR 23-JUN-1998; 98US-0090355.
 PR 24-JUN-1998; 98US-0090429.
 PR 24-JUN-1998; 98US-0090431.
 PR 24-JUN-1998; 98US-0090435.
 PR 24-JUN-1998; 98US-0090444.
 PR 24-JUN-1998; 98US-0090445.
 PR 24-JUN-1998; 98US-0090451.
 PR 24-JUN-1998; 98US-0090472.
 PR 24-JUN-1998; 98US-0090535.
 PR 24-JUN-1998; 98US-0090538.
 PR 24-JUN-1998; 98US-0090540.
 PR 25-JUN-1998; 98US-0090557.
 PR 25-JUN-1998; 98US-0090676.
 PR 25-JUN-1998; 98US-0090688.
 PR 25-JUN-1998; 98US-0090690.
 PR 25-JUN-1998; 98US-0090691.
 PR 25-JUN-1998; 98US-0090694.
 PR 25-JUN-1998; 98US-0090695.
 PR 25-JUN-1998; 98US-0090696.
 PR 26-JUN-1998; 98US-0090862.
 PR 26-JUN-1998; 98US-0090863.
 PR 01-JUL-1998; 98US-0091358.
 PR 01-JUL-1998; 98US-0091360.
 PR 01-JUL-1998; 98US-0091544.
 PR 02-JUL-1998; 98US-0091478.
 PR 02-JUL-1998; 98US-0091486.
 PR 02-JUL-1998; 98US-0091519.
 PR 02-JUL-1998; 98US-0091626.
 PR 02-JUL-1998; 98US-0091628.
 PR 02-JUL-1998; 98US-0091633.
 PR 02-JUL-1998; 98US-0091646.
 PR 02-JUL-1998; 98US-0091673.
 PR 07-JUL-1998; 98US-0091978.
 PR 09-JUL-1998; 98US-0092182.
 PR 10-JUL-1998; 98US-0092472.
 PR 20-JUL-1998; 98US-0093339.

PR 30-JUL-1998; 98US-0094651.
 PR 04-AUG-1998; 98US-0095282.
 PR 04-AUG-1998; 98US-0095285.
 PR 04-AUG-1998; 98US-0095301.
 PR 04-AUG-1998; 98US-0095302.
 PR 04-AUG-1998; 98US-0095318.
 PR 04-AUG-1998; 98US-0095321.
 PR 04-AUG-1998; 98US-0095325.
 PR 10-AUG-1998; 98US-0095916.
 PR 10-AUG-1998; 98US-0095929.
 PR 10-AUG-1998; 98US-0096012.
 PR 11-AUG-1998; 98US-0096143.
 PR 11-AUG-1998; 98US-0096146.
 PR 12-AUG-1998; 98US-0096329.
 PR 17-AUG-1998; 98US-0096757.
 PR 17-AUG-1998; 98US-0096766.
 PR 17-AUG-1998; 98US-0096768.
 PR 17-AUG-1998; 98US-0096773.
 PR 17-AUG-1998; 98US-0096791.
 PR 17-AUG-1998; 98US-0096867.
 PR 17-AUG-1998; 98US-0096891.
 PR 17-AUG-1998; 98US-0096894.
 PR 17-AUG-1998; 98US-0096895.
 PR 18-AUG-1998; 98US-0096897.
 PR 18-AUG-1998; 98US-0096949.
 PR 18-AUG-1998; 98US-0096950.
 PR 18-AUG-1998; 98US-0096959.
 PR 18-AUG-1998; 98US-0096960.
 PR 18-AUG-1998; 98US-0097022.
 PR 20-AUG-1998; 98US-0097141.
 PR 24-AUG-1998; 98US-0097218.
 PR 26-AUG-1998; 98US-0097661.
 PR 26-AUG-1998; 98US-0097951.
 PR 26-AUG-1998; 98US-0097952.
 PR 26-AUG-1998; 98US-0097954.
 PR 26-AUG-1998; 98US-0097955.
 PR 26-AUG-1998; 98US-0097971.
 PR 26-AUG-1998; 98US-0097974.
 PR 26-AUG-1998; 98US-0097978.
 PR 26-AUG-1998; 98US-0097979.
 PR 26-AUG-1998; 98US-0097986.
 PR 31-AUG-1998; 98US-0098014.
 PR 31-AUG-1998; 98US-0098525.
 PR 16-SEP-1998; 98US-0100634.
 PR 12-JAN-1999; 98US-0115565.
 XX
 PA (GETH) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK,
 PI Wood WI, Yuan J;
 WPI: 2000-072883/06.
 N-PSDB; AA264969.

Membrane-bound proteins and related nucleotide sequences -
 claim 12; Fig 47; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.

XX

SQ	Sequence	867 AA;	DT	24-OCT-2001	(first entry)
	Query Match	99.4%; Score 4723.5; DB 21; Length 867;	XX	Human PRO1120 polypeptide sequence.	
	Best Local Similarity	99.7%; Pred. No. 0;	XX	Human secretory and transmembrane; PRO; mammalian; cancer; lung;	
	Matches 867; Conservative	0; Mismatches 0; Indels 3; Gaps 1;	KW	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;	
QY	1	MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFDRNRNIRNIIILVLTDDQVBLGS 60	KW	cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;	
Db	1	MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFDRNRNIRNIIILVLTDDQVBLGS 60	XX	adipocyte; A-peptide; factor VIIA; gene therapy.	
QY	61	MOVNNKTRIMEOGGAHFINAFVTTMCCPSRSSLTGKYVHNHTYNNENCSPPSQA 120	OS	Homo sapiens.	
Db	61	MOVNNKTRIMEOGGAHFINAFVTTMCCPSRSSLTGKYVHNHTYNNENCSPPSQA 120	XX	WO200140466-A2.	
QY	121	QHSRTFAVLNSTGYRTAFKGYLNEYSVPPGKEMVGLLKNRFRNYTLCRNGVK 180	XX	07-JUN-2001.	
Db	121	QHSRTFAVLNSTGYRTAFKGYLNEYSVPPGKEMVGLLKNRFRNYTLCRNGVK 180	XX	01-DEC-2000; 2000WO-US32678.	
QY	181	EKHGSYSDVLTDLITDNDVSFFRTSKMYPHRPVLMVISHAAPHGPDAPQYSLFP 240	PR	01-DEC-1999; 99WO-US28301.	
Db	181	EKHGSYSDVLTDLITDNDVSFFRTSKMYPHRPVLMVISHAAPHGPDAPQYSLFP 240	PR	01-DEC-1999; 99WO-US28634.	
QY	241	NASQHIITSYNVAFPDKENINRYTGPKPIHMEFTNMLQKRLQTLNSVDDSMETIYNM 300	PR	02-DEC-1999; 99WO-US28551.	
Db	241	NASQHIITSYNVAFPDKENINRYTGPKPIHMEFTNMLQKRLQTLNSVDDSMETIYNM 300	PR	02-DEC-1999; 99WO-US28564.	
QY	301	LIVETGELNTIIVYTADHGYYHIGQGLVKGKSMPEYEDIRVPFVYRGNVBEAGCLNPHIV 360	PR	02-DEC-1999; 99WO-US28565.	
Db	301	LIVETGELNTIIVYTADHGYYHIGQGLVKGKSMPEYEDIRVPFVYRGNVBEAGCLNPHIV 360	PR	09-DEC-1999; 99US-0170262.	
QY	361	LNIDLAFTILDAGLDIPADMDGKILKLDTERPVNRFHLKXKRWVWDSFLVERKLL 420	PR	16-DEC-1999; 99WO-US30095.	
Db	361	LNIDLAFTILDAGLDIPADMDGKILKLDTERPVNRFHLKXKRWVWDSFLVERKLL 420	PR	20-DEC-1999; 99WO-US30911.	
QY	421	HKRDNKDVAQENFLPKYQRYKDLICQAEYQTAQCEQLGQKWCQVEDATGKLHKCKGP 480	PR	20-DEC-1999; 99WO-US30999.	
Db	421	HKRDNKDVAQENFLPKYQRYKDLICQAEYQTAQCEQLGQKWCQVEDATGKLHKCKGP 480	PR	30-DEC-1999; 99WO-US31243.	
QY	481	MELGGSRLSNLVPKYGGSEACTCDSDGYKLSLAGRRKLEFKKYKASVYRSRSIRSV 540	PR	06-JAN-2000; 2000WO-US00277.	
Db	481	MELGGSRLSNLVPKYGGSEACTCDSDGYKLSLAGRRKLEFKKYKASVYRSRSIRSV 540	PR	06-JAN-2000; 2000WO-US00376.	
QY	541	AIEVDGRVYHVLGDAAPRNLTKEHWFAGPEDDQDDKGGDFSGTGGLPDYSAANP.KVT 600	PR	11-FEB-2000; 2000WO-US03565.	
Db	541	AIEVDGRVYHVLGDAAPRNLTKEHWFAGPEDDQDDKGGDFSGTGGLPDYSAANP.KVT 600	PR	18-FEB-2000; 2000WO-US04341.	
QY	601	HRCYILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLRVGRHLKRRPEEC 660	PR	22-FEB-2000; 2000WO-US04414.	
Db	601	HRCYILENDTVQCDLDLYKSLQAWKHLDHETLQNKIKNLRVGRHLKRRPEEC 660	PR	24-FEB-2000; 2000WO-US05004.	
QY	661	DCHKISYHTQHKGRKLRGSSLHPRKGLQEKDKVWLLREOKRKKKRLKRLQNNDTC 720	PR	01-MAR-2000; 2000WO-US05601.	
Db	661	DCHKISYHTQHKGRKLRGSSLHPRKGLQEKDKVWLLREOKRKKKRLKRLQNNDTC 720	PR	20-MAR-2000; 2000WO-US07377.	
QY	721	SMPGLTCFTHDNOHQWTAFTWLTGPFCACTSANNTYWCMTINETHNPLFCFATGPLE 780	PR	21-MAR-2000; 2000WO-US07532.	
Db	721	SMPGLTCFTHDNOHQWTAFTWLTGPFCACTSANNTYWCMTINETHNPLFCFATGPLE 780	PR	30-MAR-2000; 2000WO-US08439.	
QY	781	YFDLNTDPQLMNAVNTLDRDLNQLHVLQELRSCKYKQCNPETRNMDLGLXGGSYE 840	PR	17-MAY-2000; 2000WO-US13705.	
Db	781	YFDLNTDPQLMNAVNTLDRDLNQLHVLQELRSCKYKQCNPETRNMDLGLXGGSYE 840	PR	22-MAY-2000; 2000WO-US14042.	
QY	841	QYRQFORRWKPEMKRPPSSLSGLQWEGEG 870	PR	30-MAY-2000; 2000WO-US14941.	
Db	838	QYRQFORRWKPEMKRPPSSLSGLQWEGEG 867	PR	02-JUN-2000; 2000WO-US15264.	
			XX	10-NOV-2000; 2000WO-US30873.	
			PA	(GETH) GENENTECH INC.	
			XX	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;	
			PI	Gerritsen WE, Goddard A, Godowski FJ, Gurney AL, Sherwood S;	
			PI	Smith V, Stewart TA, Tumas D, Watanabe GK, Wood WI, Zhang Z;	
			XX	WPI; 2001-408281/43.	
			DR	N-PSDB; AAS21253.	
			XX	Isolated, secretory and transmembrane PRO polypeptide used to detect	
			PT	other PRO polypeptides, link bioactive molecules to cells expressing	
			PT	PRO polypeptides, and detect the presence of mammalian tumours e.g.	
			PT	lung, breast, prostate, cervical	
			XX	Claim 12; Fig 20; 813pp; English.	
			XX	AAU12172-AAU12446 represent novel human secretory and transmembrane	
			CC	PRO polypeptides. The PRO polypeptides are useful to detect other	
			CC	PRO polypeptides, to link bioactive molecules to cells expressing	
			CC	PRO polypeptides, to modulate biological activities of cells expressing	
			CC	PRO polypeptides, and to detect the presence of mammalian lung, colon,	
			CC	breast, prostate, rectal, cervical or liver tumours by comparing PRO	
			CC	polypeptide expression in a cell sample to that in a control sample.	
			CC	Some of the 275 sequences are also useful to stimulate the release of	
			CC	tumour necrosis factor-alpha (TNF-alpha) from human blood, the	
			CC	proliferation or differentiation of chondrocytes, the proliferation or	
			CC	gene expression in pericyte cells, the release of proteoglycans from	
			CC	cartilage, the proliferation of inner ear utricular supporting cells or	

RESULT 12

AAU12181

ID AAU12181 standard; Protein; 867 AA.

XX

AC AAU12181;

XX

CC	of T-lymphocytes, the release of a cytokine from peripheral blood
CC	monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC	the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC	skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC	to factor VIIA. The PRO polypeptides can be used in assays to identify
CC	molecules involved in binding interactions. The polynucleotides encoding
CC	PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC	transgenic or knock out animals and can be used in gene therapy.
XX	
SQ	Sequence 867 AA;
Query Match 99.4%; Score 4723.5; DB 22; Length 867;	
Best Local Similarity 99.7%; Pred No. 0;	
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;	
QY	1 MGPPSLVLCISATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDQVBLGS 60
Db	1 MGPPSLVLCISATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDQVBLGS 60
QY	61 MQVMNKRIRIMEOGGAHFINAVFTPMCCPSRSSILTKYVHNNTYNNENCSSPSWQA 120
Db	61 MQVMNKRIRIMEOGGAHFINAVFTPMCCPSRSSILTKYVHNNTYNNENCSSPSWQA 120
QY	121 QHESRTFAVYLNSTGYRTAFPGKLYNEVSGYVPPGKEMVGLLKNSRFNYTLCRNGVK 180
Db	121 QHESRTFAVYLNSTGYRTAFPGKLYNEVSGYVPPGKEMVGLLKNSRFNYTLCRNGVK 180
QY	181 EKHGSDYSKDYLTDLITNDSVFFRTSKMYPHPVLMVISHAAPHGPDSPAYSRLFP 240
Db	181 EKHGSDYSKDYLTDLITNDSVFFRTSKMYPHPVLMVISHAAPHGPDSPAYSRLFP 240
QY	241 NASQHIITSYNAFNPDKHMTVYTPGKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db	241 NASQHIITSYNAFNPDKHMTVYTPGKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY	301 LVEGELDNITVIYVTDAGHYHIGFGLVKGSMYEFDIRVPYVRGPNVEAGCLNPHIV 360
Db	301 LVEGELDNITVIYVTDAGHYHIGFGLVKGSMYEFDIRVPYVRGPNVEAGCLNPHIV 360
QY	361 LNIDLAPILDIAGLDIPADMDGKSILKLDTERPVRNREHLKKRWYRDSFLVERKLL 420
Db	361 LNIDLAPILDIAGLDIPADMDGKSILKLDTERPVRNREHLKKRWYRDSFLVERKLL 420
QY	421 HKRDNKDYDAEENELPKYQRYKDLQRAEYQYACELGQKQWCVEDATGKLHKCKG 480
Db	421 HKRDNKDYDAEENELPKYQRYKDLQRAEYQYACELGQKQWCVEDATGKLHKCKG 480
QY	481 MRLGSRALSNLVPKYCGGSEACTCDGDKYLSLGERKLLFKKKYKASYRVSRSRSV 540
Db	481 MRLGSRALSNLVPKYCGGSEACTCDGDKYLSLGERKLLFKKKYKASYRVSRSRSV 540
QY	541 ALEVDRVYHVGLGDAQFRLNLTKRHWFGAPEDQDDGDFGSGTGLPDYSAANPKVT 600
Db	541 ALEVDRVYHVGLGDAQFRLNLTKRHWFGAPEDQDDGDFGSGTGLPDYSAANPKVT 600
QY	601 HRCYILENDTVQCDLDLYKSLQANKDKHLHDHEIETLQNKIKNLRVGRHLKKRPEEC 660
Db	601 HRCYILENDTVQCDLDLYKSLQANKDKHLHDHEIETLQNKIKNLRVGRHLKKRPEEC 660
QY	661 DCHKLSYHTQHKGRHLKRGSSILHPPKGLQEKDKVWLLREOKRKKLKLRLQNNDC 720
Db	661 DCHKLSYHTQHKGRHLKRGSSILHPPKGLQEKDKVWLLREOKRKKLKLRLQNNDC 720
QY	721 SMPGLTCTHNDQHWQTAFTWTLGPFCACTSANNTYVCMRTINETNFIKCEPATGFLE 780
Db	721 SMPGLTCTHNDQHWQTAFTWTLGPFCACTSANNTYVCMRTINETNFIKCEPATGFLE 780
QY	781 YFDLNTDYPQLMNAVNTLDRVLNQLHVLMELSCKYKOCNPRTRMDLGLKDGGSYE 840
Db	781 YFDLNTDYPQLMNAVNTLDRVLNQLHVLMELSCKYKOCNPRTRMDLGLKDGGSYE 837
QY	841 QYRQFQRRKWPMPKRPSSKSLQLMWEGWEG 870

Db	838 QYRQFQRRKWPMPKRPSSKSLQLMWEGWEG 867
RESULT 13	
AA865171	
ID	AAB65171 standard; Protein; 867 AA.
XX	AC AAB65171;
XX	02-APR-2001 (first entry)
XX	Human PRO1120 (UNQ559) protein sequence SEQ ID NO:84.
KW	Human; secreted and transmembrane protein; PRO; cytosstatic;
KW	cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
XX	diagnostic assay.
OS	Homo sapiens.
PN	WO200073454-A1.
PD	07-DEC-2000.
XX	30-MAR-2000; 2000WO-US08439.
PR	02-JUN-1999; 99WO-US12252.
PR	23-JUN-1999; 99US-0141037.
PR	07-JUL-1999; 99US-0143048.
PR	20-JUL-1999; 99US-0144758.
PR	26-JUL-1999; 99US-0145698.
PR	28-JUL-1999; 99US-0146222.
PR	17-AUG-1999; 99US-0149396.
PR	15-SEP-1999; 99WO-US21090.
PR	15-SEP-1999; 99WO-US21547.
PR	08-OCT-1999; 99US-0158663.
PR	30-NOV-1999; 99WO-US28313.
PR	01-DEC-1999; 99WO-US28301.
PR	16-DEC-1999; 99WO-US30095.
PR	20-DEC-1999; 99WO-US30911.
PR	05-JAN-2000; 2000WO-US00219.
PR	06-JAN-2000; 2000WO-US00376.
PR	11-FEB-2000; 2000WO-US03565.
PR	18-FEB-2000; 2000WO-US04341.
PR	22-FEB-2000; 2000WO-US04414.
PR	24-FEB-2000; 2000WO-US04914.
PR	02-MAR-2000; 2000WO-US05004.
PR	02-MAR-2000; 2000WO-US05841.
PR	15-MAR-2000; 2000WO-US06884.
PR	20-MAR-2000; 2000WO-US07377.
XX	(GETH) GENENTECH INC.
XX	Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI	Perrara N, Pong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI	Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
PI	Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
XX	Zhang Z;
XX	WPI; 2001-032160/04.
DR	N-PSDB; AAF44115.
XX	
PT	PRO polynucleotides used to produce polypeptides used to target
PT	bioactive molecules such as toxins, radiolabels or antibodies, to
PT	specific cells, to cause targeted cell death -
XX	
PS	Claim 12; Fig 47; 935pp; English.
XX	
CC	The present invention describes human secreted and transmembrane PRO
CC	proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC	can be used for targeted delivery of bioactive molecules, such as
CC	toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC	sequences, and their fragments, can be used as hybridisation probes, in
CC	chromosomal and gene mapping, and in the generation of anti-sense RNA

CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
SQ Sequence 867 AA;

Query Match 99.4%; Score 4723.5; DB 22; Length 867;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MGPPSLVLCILSNVPSILGGSFAFLSHRLKGRFQDRNRIRPMILVLDDQVLEGS 60
Db |||||
1 MGPPSLVLCILSNVPSILGGSFAFLSHRLKGRFQDRNRIRPMILVLDDQVLEGS 60
QY 61 MQVMNKTREIMEQGAHFINAFVTPMCCPSRSSILTKYVHNHTYNNENCSPPSQA 120
Db |||||
61 MQVMNKTREIMEQGAHFINAFVTPMCCPSRSSILTKYVHNHTYNNENCSPPSQA 120
QY 121 QHESRTFAVILNSTGTRTAFKYLNEYNGSVTPGKWEKWLKNSRPNYTLCRNGVK 180
Db |||||
121 QHESRTFAVILNSTGTRTAFKYLNEYNGSVTPGKWEKWLKNSRPNYTLCRNGVK 180
QY 181 EXHGSYSKDYLTLTNDVSFFRTSKWYPHRPVLWVISHAAPHGSDAPQVSRFLP 240
Db |||||
181 EXHGSYSKDYLTLTNDVSFFRTSKWYPHRPVLWVISHAAPHGSDAPQVSRFLP 240
QY 241 NASQHITPSYNAPNDPKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDSMETIYNN 300
Db |||||
241 NASQHITPSYNAPNDPKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDSMETIYNN 300
QY 301 LYETGELNTLYVTADHGYHIGQGLVKGSMPYEFDIRVPFYVGRGVNVEAGCLNPHIV 360
Db |||||
301 LYETGELNTLYVTADHGYHIGQGLVKGSMPYEFDIRVPFYVGRGVNVEAGCLNPHIV 360
QY 361 LNIDLAFTILDLAGLIDIPADMDGKSLKLLDTERPVNRFHLKKQVWRDVSFLVERGKLL 420
Db |||||
361 LNIDLAFTILDLAGLIDIPADMDGKSLKLLDTERPVNRFHLKKQVWRDVSFLVERGKLL 420
QY 421 HKRDNDKVAQENELPKYQVKDLCQAEYQACEQLGQKWCQVEDATGKLHKCKGP 480
Db |||||
421 HKRDNDKVAQENELPKYQVKDLCQAEYQACEQLGQKWCQVEDATGKLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKCLKKXKASVYGRSIRSV 540
Db |||||
481 MRLGGSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKCLKKXKASVYGRSIRSV 540
QY 541 ALEVGRVHVGLGDAAPRNLTQRHWPAGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600
Db |||||
541 ALEVGRVHVGLGDAAPRNLTQRHWPAGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600
QY 601 HRCYILENTVQCDLLDYLSLQAWDKLHIDHETLQNKIKNREVRGHLKKRPEEC 660
Db |||||
601 HRCYILENTVQCDLLDYLSLQAWDKLHIDHETLQNKIKNREVRGHLKKRPEEC 660
QY 661 DCHKISYTHQKRLKRGSSLHPFRKGLQEKDKYVLLREQKRLKRLKRLQNDTC 720
Db |||||
661 DCHKISYTHQKRLKRGSSLHPFRKGLQEKDKYVLLREQKRLKRLKRLQNDTC 720
QY 721 SMPLGLTCFTHDNOHWQAPFWTLGPFCACTSANNTYWCMTINETHNFLECBATGPLE 780
Db |||||
721 SMPLGLTCFTHDNOHWQAPFWTLGPFCACTSANNTYWCMTINETHNFLECBATGPLE 780
QY 781 YFDLNTDPPQLNNAVNTLDRVNLQHLVQLMELRSCKYKQCNPTRNMDLGLDGGSYE 840
Db |||||
781 YFDLNTDPPQLNNAVNTLDRVNLQHLVQLMELRSCKYKQCNPTRNMDLGLDGGSYE 840
QY 841 QYRQFORRWPEKRPSSKSLGQLWEGWEG 870
Db |||||

Db 838 QYRQFORRWPEKRPSSKSLGQLWEGWEG 867

RESULT 14

ID ABB95459 standard; Protein; 867 AA.

AC ABB95459;

XX 19-JUL-2002 (first entry)

DE Human angiogenesis related protein PRO1120 SEQ ID NO: 74.

KW Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;
KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;
KW cardiant; cytostatic; antiangiogenic; hypotensive; vulnerary;
XX antiarteriosclerotic.

OS Homo sapiens.

XX WO200208284-A2.

XX 31-JAN-2002.

XX 09-JUL-2001; 2001WO-US21735.

XX 20-JUL-2000; 2000US-219556P.

XX 25-JUL-2000; 2000US-220624P.

XX 25-JUL-2000; 2000US-220684P.

XX 28-JUL-2000; 2000WO-US20710.

XX 02-AUG-2000; 2000US-222695P.

XX 17-AUG-2000; 2000US-0643657.

XX 23-AUG-2000; 2000WO-US23322.

XX 24-AUG-2000; 2000WO-US23328.

XX 07-SEP-2000; 2000US-230978P.

XX 15-SEP-2000; 2000US-000000P.

XX 18-SEP-2000; 2000US-0664610.

XX 18-SEP-2000; 2000US-0665350.

XX 08-OCT-2000; 2000US-242922P.

XX 08-NOV-2000; 2000US-0709238.

XX 10-NOV-2000; 2000WO-US30873.

XX 01-DEC-2000; 2000WO-US32878.

XX 20-DEC-2000; 2000US-0747259.

XX 22-DEC-2000; 2000WO-US34956.

XX 22-JAN-2001; 2001US-0767609.

XX 28-FEB-2001; 2001US-0796498.

XX 01-MAR-2001; 2001WO-US06520.

XX 09-MAR-2001; 2001US-0802706.

XX 14-MAR-2001; 2001US-0808689.

XX 22-MAR-2001; 2001US-0816744.

XX 05-APR-2001; 2001US-0828366.

XX 10-MAY-2001; 2001US-0854208.

XX 10-MAY-2001; 2001US-0854280.

XX 25-MAY-2001; 2001US-0866028.

XX 25-MAY-2001; 2001US-0866034.

XX 25-MAY-2001; 2001WO-US17092.

XX 30-MAY-2001; 2001US-0870574.

XX 01-JUN-2001; 2001WO-US17443.

XX 20-JUN-2001; 2001WO-US19692.

XX 28-JUN-2001; 2001WO-US00000.

(GETH) GENENTECH INC.

(BAKE/) BAKER K P.

(FERR/) FERRARA N.

(GERB/) GERBER H.

(GERR/) GERRITSEN M E.

(GODD/) GODDARD A.

(GODO/) GODOWSKI P J.

(GURN/) GURNEY A L.

(HILL/) HILLAN K J.

PR 01-MAR-2001; 2001WO-US06566.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0808689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 28-MAY-2001; 2001US-0866028.
 PR 25-MAY-2001; 2001US-0866034.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 30-MAY-2001; 2001US-0870574.
 PR 30-MAY-2001; 2001WO-US17443.
 PR 01-JUN-2001; 2001WO-US17800.

(GETH) GENENTECH INC.

PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;
 PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;
 PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;

XX WPI: 2002-090516/12.
 DR N-PSDB; ABL88108.

XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,
 PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial
 PT infarction), endothelial or angiogenic disorders in a mammal -
 XX Claim 11; Fig 74; 565pp; English.

XX ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to
 CC ABB85003. The PRO proteins and polynucleotides have cardiant, cytostatic,
 CC antiangiogenic, hypotensive, vulnerary and antiarteriosclerotic
 CC activities, and can be used in gene therapy. The PRO polynucleotides,
 CC proteins, agonists and antagonists are useful for treating or diagnosing
 CC a cardiovascular, endothelial or angiogenic disorder in a mammal,
 CC e.g. cardiac hypertrophy, trauma, cancer, age-related macular
 CC degeneration, atherosclerosis, hypertension, arterial restenosis,
 CC rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis,
 CC lymphangitis, tumour angiogenesis (such as breast carcinoma and liver
 CC carcinoma) and wound healing. The PRO polynucleotides have applications
 CC in molecular biology, including use as hybridisation probes, and in
 CC chromosome and gene mapping. ABL88259 to ABL88267 represent primers and
 CC probes used in the exemplification of the present invention.

XX Sequence 867 AA;

Query Match 99.4%; Score 4723.5; DB 23; Length 867;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
 QY 1 MGPPSLVCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60
 Db 1 MGPPSLVCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60
 QY 61 MOVNKRTRIMEGGGAHFNAFTVTMCCPSRSSILTKGYVHNHTYNNENCSPPSQA 120
 Db 61 MOVNKRTRIMEGGGAHFNAFTVTMCCPSRSSILTKGYVHNHTYNNENCSPPSQA 120
 QY 121 QHESRTFAVYLNSTGYRTAFGKYLNEYNGSVPPGKWEVGLLKNRFRNYNLTLCRNGVK 180
 Db 121 QHESRTFAVYLNSTGYRTAFGKYLNEYNGSVPPGKWEVGLLKNRFRNYNLTLCRNGVK 180
 QY 181 EKHGSDYSKDYLTDLITNDSVFPFTSKMGPHRVLVMI SHAAPHGPEDSAPQYSRLFP 240
 Db 181 EKHGSDYSKDYLTDLITNDSVFPFTSKMGPHRVLVMI SHAAPHGPEDSAPQYSRLFP 240
 QY 241 NASQHITPSYNAFPNDKHVIMRYTGPMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNN 300
 Db 241 NASQHITPSYNAFPNDKHVIMRYTGPMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNN 300
 QY 301 LVETGELDNITYIVYTADHGHIHQGLVKGKSMYPYFDIRVPFYVGRPNVEAGCLNPHIV 360
 Db 301 LVETGELDNITYIVYTADHGHIHQGLVKGKSMYPYFDIRVPFYVGRPNVEAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLL 420
 Db 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLL 420
 QY 421 HKDNDKVDQAQENFLPKYQVRVKDLCORABYOTACEQLGQKWCQVEDATGKLHKCKGP 480
 Db 421 HKDNDKVDQAQENFLPKYQVRVKDLCORABYOTACEQLGQKWCQVEDATGKLHKCKGP 480
 QY 481 MRLGGSRLSNLVPKYVGQSEACTCDSDYKLSLAGRRKLFKKYKASYVRSRIRSV 540
 Db 481 MRLGGSRLSNLVPKYVGQSEACTCDSDYKLSLAGRRKLFKKYKASYVRSRIRSV 540
 QY 541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 Db 541 AIEVDGRVHVGLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 QY 601 HRCVILENDTVQCDLDLYKSLQAKDKHLHIDHETLQNKIKNLRVEVGHLLKKRPEEC 660
 Db 601 HRCVILENDTVQCDLDLYKSLQAKDKHLHIDHETLQNKIKNLRVEVGHLLKKRPEEC 660
 QY 721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNTYWCMTINETHNLFCEFATGFLE 780
 Db 721 SMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNTYWCMTINETHNLFCEFATGFLE 780
 QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNNMDLGLKDGGSYE 840
 Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTNNMDLGLKDGGSYE 840
 QY 841 QYRQFORRKWPEMKRPSKSLGQWEGWEG 870
 Db 838 QYRQFORRKWPEMKRPSKSLGQWEGWEG 867

Search completed: February 14, 2004, 22:54:05
 Job time : 76 secs

1	3363	100.0	3363	13	US-10-025-966A-4	Sequence 4
2	3363	100.0	3363	13	US-10-265-071-4	Sequence 4
3	3348.4	99.6	4303	12	US-10-264-237-1349	Sequence 1
4	3347.2	99.5	4397	13	US-09-970-287-11	Sequence 1
5	3346.8	99.5	4286	13	US-10-025-966A-13	Sequence 1
6	3346.8	99.5	4286	13	US-10-265-071-13	Sequence 1
7	3344.8	99.5	3348	13	US-10-183-951-17	Sequence 1
8	3323.4	98.8	3306	9	US-09-989-722-83	Sequence 8
9	3323.4	98.8	3306	9	US-09-989-722-83	Sequence 8
10	3323.4	98.8	3306	9	US-09-989-722-83	Sequence 8
11	3323.4	98.8	3306	9	US-09-989-722-83	Sequence 8
12	3323.4	98.8	3306	10	US-09-989-722-83	Sequence 8
13	3323.4	98.8	3306	10	US-09-989-722-83	Sequence 8
14	3323.4	98.8	3306	10	US-09-991-073-83	Sequence 8
15	3323.4	98.8	3306	10	US-09-990-442-83	Sequence 8

	Query Match	100.0%;	Score 3363;	DB 13;	Length 3363;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3363;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	GACTCCCGCATCCCAAAAGACACAGATCAGCAAAAAAAGATGGGCCCCCGGAGC	60		
Db	1	GACTCCCGCATCCCAAAAGACACAGATCAGCAAAAAAAGATGGGCCCCCGGAGC	60		
QY	61	CTCGTGCTGTGCTGTGTCGCGAACTGTGTTCTCCCTGCTGGGTGGAGCTCGGCCTTC	120		
Db	61	CTCGTGCTGTGCTGTGTCGCGAACTGTGTTCTCCCTGCTGGGTGGAGCTCGGCCTTC	120		
QY	121	CTGTGCGACCAACCGCTGAAAGCGCAGGTTTCAGAGSGACCGCAGGAAACATTCGCGCCCCAAC	180		

121 CTGTGCGCACCGCGCTGAAAGGCGAGGTTTCAGAGGAGCCGAGGAAATCCGCGCCCAAC 180
181 ATCATCTCTGGTGTCTACGAGACGACAGATGTGGAGCTGGGTTCCATGCGAGGTATGAAC 240
181 ATCATCTCTGGTGTCTACGAGACGACAGATGTGGAGCTGGGTTCCATGCGAGGTATGAAC 240
241 AAGACCGCGCGCATCATGGAGCAGCGCGCGGCGCACTTCATCAACGCTTCGTGACCAACA 300
241 AAGACCGCGCGCATCATGGAGCAGCGCGCGGCGCACTTCATCAACGCTTCGTGACCAACA 300
301 CCCATGTGTGCGCCCTCAGCTCTCCATCTCTCACTGSCAAGTACGTCCCAACCAAC 360
301 CCCATGTGTGCGCCCTCAGCTCTCCATCTCTCACTGSCAAGTACGTCCCAACCAAC 360
361 ACCTACACCAATGAGAATGCTCTCTCGCCCTCTCTGCGGACAGCAGCAGAGCGCGC 420
361 ACCTACACCAATGAGAATGCTCTCTCGCCCTCTCTGCGGACAGCAGCAGAGCGCGC 420
421 ACCTTGGCGTGTACTCAATAGCACTGGCTACCGGACAGCTTTCTTGGGAAGTATCTT 480
421 ACCTTGGCGTGTACTCAATAGCACTGGCTACCGGACAGCTTTCTTGGGAAGTATCTT 480
481 AATGAATACACCGCTCTCACTGTCACCGCGCTGGAAGAGTGGTCCGACTCCTTAAA 540
481 AATGAATACACCGCTCTCACTGTCACCGCGCTGGAAGAGTGGTCCGACTCCTTAAA 540
541 AACTCCGCTTTTAACTAATACAGCTGTGTGGAACCGGGGTGAAGAAAGCAGCGTCC 600
541 AACTCCGCTTTTAACTAATACAGCTGTGTGGAACCGGGGTGAAGAAAGCAGCGTCC 600
601 GACTACTCCAAAGATTAACCTACAGACCTCATCAACCAATGACAGGTAGCTTCTCCG 660
601 GACTACTCCAAAGATTAACCTACAGACCTCATCAACCAATGACAGGTAGCTTCTCCG 660
661 AGCTCCAAAGATTAACCTACAGACCTCATCAACCAATGACAGGTAGCTTCTCCG 720
661 AGCTCCAAAGATTAACCTACAGACCTCATCAACCAATGACAGGTAGCTTCTCCG 720
721 CAGGCGCTGAGGATTCAGCCCCCAATATTCAGCGCTCTTCCCAAGCGATCTCAGCAC 780
721 CAGGCGCTGAGGATTCAGCCCCCAATATTCAGCGCTCTTCCCAAGCGATCTCAGCAC 780
781 ATCAGCGCGAGCTCAACTACCGCGCCCAACCGGACAACTGGATCATCGGTACACG 840
781 ATCAGCGCGAGCTCAACTACCGCGCCCAACCGGACAACTGGATCATCGGTACACG 840
841 GGGCCCATGAAGCCCATCCATGGAATTCACCAATGCTCCAGCGGAAGCGCTTGCGAG 900
841 GGGCCCATGAAGCCCATCCATGGAATTCACCAATGCTCCAGCGGAAGCGCTTGCGAG 900
901 ACCCTCATGTGCGTGGACGACTCCATGGAGAGATTTACACATGCTGGTTGAGACGGGC 960
901 ACCCTCATGTGCGTGGACGACTCCATGGAGAGATTTACACATGCTGGTTGAGACGGGC 960
961 GAGCTGGACAAACGTAATGATATACCGCGGACCAAGGTTTACCAATGCTGGTTGAGACGGGC 1020
961 GAGCTGGACAAACGTAATGATATACCGCGGACCAAGGTTTACCAATGCTGGTTGAGACGGGC 1020
1021 GGCCTGGTGAAGGGAATTCATGCTATGATGTTGATCATGAGGTCCCGTTCTACGTG 1080
1021 GGCCTGGTGAAGGGAATTCATGCTATGATGTTGATCATGAGGTCCCGTTCTACGTG 1080
1081 AGGGGCCCCAAGCTGGAAGCGGCTGTCTGAATCCCCACATGCTCCTCAACATTTGACCTG 1140
1081 AGGGGCCCCAAGCTGGAAGCGGCTGTCTGAATCCCCACATGCTCCTCAACATTTGACCTG 1140
1141 GCGCCCAACATCTGGAACATTCAGCGCTTGGACATACCTGCGGATATGGACGGGAATCC 1200
1141 GCGCCCAACATCTGGAACATTCAGCGCTTGGACATACCTGCGGATATGGACGGGAATCC 1200
1201 ATCTCTCAAGCTGTGGACACGAGCGCGCGGTGAATCGGTTTCACTTCAAAAAAGAGATG 1260

1201 ATCTCTCAAGCTGTGGACACGAGCGCGCGGTGAATCGGTTTCACTTGAAAAAGAGATG 1260
1261 AGGGTCTGCGGGAATCTCTTTTGTGGAGAGAGCAAGCTCTACACAAGAGAGACAAT 1320
1261 AGGGTCTGCGGGAATCTCTTTTGTGGAGAGAGCAAGCTCTACACAAGAGAGACAAT 1320
1321 GACAAGGTGGACGCGGAGAGAACTTTTCTGCCAAGTACCAAGCTGTGAAGACCTG 1380
1321 GACAAGGTGGACGCGGAGAGAACTTTTCTGCCAAGTACCAAGCTGTGAAGACCTG 1380
1381 TGTACGCTGCTGAGTACCAAGACGCGGTGTGAGAGCTGGGACAGAAGTGGCAAGTGTG 1440
1381 TGTACGCTGCTGAGTACCAAGACGCGGTGTGAGAGCTGGGACAGAAGTGGCAAGTGTG 1440
1441 GAGGACGCCACCGGGAAGCTGAAGCTCAAGTCAAGGGGCCCATCGGGTGGCGGCG 1500
1441 GAGGACGCCACCGGGAAGCTGAAGCTCAAGTCAAGGGGCCCATCGGGTGGCGGCG 1500
1501 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGGGCAGGGCAGGAGCCTGCACC 1560
1501 AGCAGAGCCCTCTCCAACTCTGTCGCCCAAGTACTAGGGCAGGGCAGGAGCCTGCACC 1560
1561 TGTGACAGCGGGAATCAAGCTCAGCTGGCGGACCGCGGAAAAAACTCTTCAAGAG 1620
1561 TGTGACAGCGGGAATCAAGCTCAGCTGGCGGACCGCGGAAAAAACTCTTCAAGAG 1620
1621 AAGTACAGCGCAGCTATGTCCGAGTCCCTCATCCGCTCAGTGGCCATCGAGTGGAC 1680
1621 AAGTACAGCGCAGCTATGTCCGAGTCCCTCATCCGCTCAGTGGCCATCGAGTGGAC 1680
1681 GGCGAGGCTGTACACGTAGGCTGTGATGCGCGCCAGCGCCGAAAAAACTCTCAAGGG 1740
1681 GGCGAGGCTGTACACGTAGGCTGTGATGCGCGCCAGCGCCGAAAAAACTCTCAAGGG 1740
1741 CACTGCGCAGGGGCCCCGTGAGCAACAGATGACAAAGATGGTGGGACTTCACTGGCACT 1800
1741 CACTGCGCAGGGGCCCCGTGAGCAACAGATGACAAAGATGGTGGGACTTCACTGGCACT 1800
1801 GGAGGCTTCCGAGTACTCAGCGCCCAACCCATTAAAGTGACACATCGGTGCTACATC 1860
1801 GGAGGCTTCCGAGTACTCAGCGCCCAACCCATTAAAGTGACACATCGGTGCTACATC 1860
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGCAGGCTGGAAA 1920
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGCAGGCTGGAAA 1920
1921 GACCAACAGCTGCAATCGACCAAGATTTGAAACCTCTGCAGAACAAATTAAGAACCTG 1980
1921 GACCAACAGCTGCAATCGACCAAGATTTGAAACCTCTGCAGAACAAATTAAGAACCTG 1980
1981 AGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCCAGAGATGTGACTGTCACAAAATC 2040
1981 AGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCCAGAGATGTGACTGTCACAAAATC 2040
2041 AGTACCAACCGCAGCAAAAGCGGCTCTCAAGCAAGAGGCTCAGTCTGATCTCTTTC 2100
2041 AGTACCAACCGCAGCAAAAGCGGCTCTCAAGCAAGAGGCTCAGTCTGATCTCTTTC 2100
2101 AGGAGGGCTCTGAGAGAGAGGACAGAGTGTGCTGTTGCGGGAGCAGAGCGCAAGAG 2160
2101 AGGAGGGCTCTGAGAGAGAGGACAGAGTGTGCTGTTGCGGGAGCAGAGCGCAAGAG 2160
2161 AAACTCCGCAAGCTGCTCAAGCGCTTCAGGAAACCAACGACACGCTGAGCATGCGAGGCTC 2220
2161 AAACTCCGCAAGCTGCTCAAGCGCTTCAGGAAACCAACGACACGCTGAGCATGCGAGGCTC 2220
2221 AGTGTCTTCAACCAAGCAACGAGCTGGGAGCGGCTTTCTTGACACTGCGGCT 2280
2221 AGTGTCTTCAACCAAGCAACGAGCTGGGAGCGGCTTTCTTGACACTGCGGCT 2280
2281 TTCTGTGCTGACACGAGCGCCCAATTAACAGTACTTGGTGCATGAGGACCAATCAATGAG 2340
2281 TTCTGTGCTGACACGAGCGCCCAATTAACAGTACTTGGTGCATGAGGACCAATCAATGAG 2340

```
QY 2341 ACTCAAAATTTCTCTCTCTCTGTAATTTGCAACTGCTTCTTAGAGTACTTTGATCTCAAC 2400
Db |||||
QY 2341 ACTCAAAATTTCTCTCTCTGTAATTTGCAACTGCTTCTTAGAGTACTTTGATCTCAAC 2400
Db |||||
QY 2401 ACAGACCCCTACCAGCTGATGATGAGTGAACACACTGGACAGGATGCTCTCAACCG 2460
Db |||||
QY 2401 ACAGACCCCTACCAGCTGATGATGAGTGAACACACTGGACAGGATGCTCTCAACCG 2460
Db |||||
QY 2461 CTACAGCTACAGCTCATGAGCTGAGAGCTCAAGGTTTCAAGAGTGAACCCCGG 2520
Db |||||
QY 2461 CTACAGCTACAGCTCATGAGCTGAGAGCTCAAGGTTTCAAGAGTGAACCCCGG 2520
Db |||||
QY 2521 ACTCGAAACATGAGCTGAGCTTAAAGATGAGAGTGAACCTTCTTCCAAATCACTGGGACACTGTGG 2580
Db |||||
QY 2521 ACTCGAAACATGAGCTGAGCTTAAAGATGAGAGTGAACCTTCTTCCAAATCACTGGGACACTGTGG 2580
Db |||||
QY 2581 CAGCTCGAAAGTGCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640
Db |||||
QY 2581 CAGCTCGAAAGTGCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640
Db |||||
QY 2641 GAAGCTCGGAGGTTAAGAACACAGAGTGGACCTCCAAACATAGAGGATCAAC 2700
Db |||||
QY 2641 GAAGCTCGGAGGTTAAGAACACAGAGTGGACCTCCAAACATAGAGGATCAAC 2700
Db |||||
QY 2701 TGACTGCACAGCAATGAAACACCATGTGGTGAATTTCCAGCAGACCTGTCTATTGGCC 2760
Db |||||
QY 2701 TGACTGCACAGCAATGAAACACCATGTGGTGAATTTCCAGCAGACCTGTCTATTGGCC 2760
Db |||||
QY 2761 AGGAGCTCGAAGACGACGACCTCTCAGTCAACATGACAGATTTGGAGGATAAC 2820
Db |||||
QY 2761 AGGAGCTCGAAGACGACGACCTCTCAGTCAACATGACAGATTTGGAGGATAAC 2820
Db |||||
QY 2821 CAGCAGGACGAGATACTTTCAGAGTCCATTTTGGCCCTCTCTTTGCTTTGGATTA 2880
Db |||||
QY 2821 CAGCAGGACGAGATACTTTCAGAGTCCATTTTGGCCCTCTCTTTGCTTTGGATTA 2880
Db |||||
QY 2881 TACCTCACAGCTGCACAAAATGATTTTTCGATCAAAAGTCAACACTAACCCCTCCC 2940
Db |||||
QY 2881 TACCTCACAGCTGCACAAAATGATTTTTCGATCAAAAGTCAACACTAACCCCTCCC 2940
Db |||||
QY 2941 CCAGAGCTCAAGAGGAAACGAGAGAGGAGAGAGATTTCTTGGAAATTTCT 3000
Db |||||
QY 2941 CCAGAGCTCAAGAGGAAACGAGAGAGGAGAGAGATTTCTTGGAAATTTCT 3000
Db |||||
QY 3001 CCCAAGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCAGTCTCTGTTCTAAA 3060
Db |||||
QY 3001 CCCAAGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCAGTCTCTGTTCTAAA 3060
Db |||||
QY 3061 TCCTCTTATTCTTTGTTGTTGTCACAAAGAGGAACTAAGAGCAGACAGAGCAACGT 3120
Db |||||
QY 3061 TCCTCTTATTCTTTGTTGTTGTCACAAAGAGGAACTAAGAGCAGACAGAGCAACGT 3120
Db |||||
QY 3121 GGAGAGGCTGAAAACAGTCAGACAGCTTTGACAAATGAGTCAGTAGCACAAGAGAGATGA 3180
Db |||||
QY 3121 GGAGAGGCTGAAAACAGTCAGACAGCTTTGACAAATGAGTCAGTAGCACAAGAGAGATGA 3180
Db |||||
QY 3181 CATTTACTAGCATTATAACCCCTGTTGCTCTGAAAGAACTCCCTTCATTGTATATAT 3240
Db |||||
QY 3181 CATTTACTAGCATTATAACCCCTGTTGCTCTGAAAGAACTCCCTTCATTGTATATAT 3240
Db |||||
QY 3241 GTGACTATTATACATGTAATCAACATGGAATTTTAGGGAACTTAATAGAAATCCCA 3300
Db |||||
QY 3241 GTGACTATTATACATGTAATCAACATGGAATTTTAGGGAACTTAATAGAAATCCCA 3300
Db |||||
QY 3301 TTTTCAGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAAA 3360
Db |||||
QY 3301 TTTTCAGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAAA 3360
Db |||||
QY 3361 AAA 3363
Db |||||
QY 3361 AAA 3363
Db |||||
```

```
RESULT 2
US-10-265-071-4
; Sequence 4, Application US/10265071
; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3363
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-265-071-4

Query Match 100.0%; Score 3363; DB 13; Length 3363;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACTCCCGCATCCCCAAAGAACACACAGATCAGCAAAAAAAGAAAGATGGCCCCCGAGC 60
Db 1 GACTCCCGCATCCCCAAAGAACACACAGATCAGCAAAAAAAGAAAGATGGCCCCCGAGC 60
QY 61 CTGCTGCTGTGCTGCTGTGCTGCGCAACTGTGTTCTCCCTGCTGGTGGAGCTCGCCCTTC 120
Db 61 CTGCTGCTGTGCTGCTGTGCTGCGCAACTGTGTTCTCCCTGCTGGTGGAGCTCGCCCTTC 120
QY 121 CTGTCGCACACCGCTGAAAGGAGGTTTCAGAGGACCGGAGGAAATCGGCCCAAC 180
Db 121 CTGTCGCACACCGCTGAAAGGAGGTTTCAGAGGACCGGAGGAAATCGGCCCAAC 180
QY 181 ATCATCTCTGGTCTCAGCGACACACAGATGTGGAGCTGGGTTCATCAGGTGATGAAC 240
Db 181 ATCATCTCTGGTCTCAGCGACACACAGATGTGGAGCTGGGTTCATCAGGTGATGAAC 240
QY 241 AAGACCCCGCGCATCATGAGAGCGGGCGGCGCACTTCATCAACGCCTTCGTGACCA 300
Db 241 AAGACCCCGCGCATCATGAGAGCGGGCGGCGCACTTCATCAACGCCTTCGTGACCA 300
QY 301 CCCATGTGCTGCCCTCCTCAGCTCCTCCATCTCTGCAAGTAGTCCCAACACCAAC 360
Db 301 CCCATGTGCTGCCCTCCTCAGCTCCTCCATCTCTGCAAGTAGTCCCAACACCAAC 360
QY 361 ACCTACACCAAAATGAAACTGTCTCTCGCCCTCTGCGAGGACAGCAGAGAGCGCG 420
Db 361 ACCTACACCAAAATGAAACTGTCTCTCGCCCTCTGCGAGGACAGCAGAGAGCGCG 420
QY 421 ACCTTTGCGGTGTAACCTCAATAGCACTGGCTACCGGACAGCTTTCTCGGAAGTATCTT 480
Db 421 ACCTTTGCGGTGTAACCTCAATAGCACTGGCTACCGGACAGCTTTCTCGGAAGTATCTT 480
QY 481 AATGAATACACGGCTCTCTAGTCCACCCCGGCTGGAGGAGTGGGTCTGACTCTTTAAA 540
Db 481 AATGAATACACGGCTCTCTAGTCCACCCCGGCTGGAGGAGTGGGTCTGACTCTTTAAA 540
QY 541 AACTCCCGCTTTTATTAATACTACGCTGTGTCGAAACGGGGTGAAGAAAGACAGCGCTCC 600
Db 541 AACTCCCGCTTTTATTAATACTACGCTGTGTCGAAACGGGGTGAAGAAAGACAGCGCTCC 600
QY 601 GACTACTCCAAGGATTACCTCACAGACCTCATCAATCAATGACAGCTGAGCTTCTTCGC 660
Db |||||
```


Db 601 GACTACTCCAGGATTACCTCAGACCTCATACCAATGACAGCGTAGCTTTCCCGC 660
Qy 661 ACCTCAAGAGATGTATCCCGCA CAGCGCAAGTCTCTCATGTCTATCAGCCATGACGCCCC 720
Db 661 ACCTCAAGAGATGTATCCCGCA CAGCGCAAGTCTCTCATGTCTATCAGCCATGACGCCCC 720
Qy 721 CACGGCCTGAGGATTGAGCCCA CCAATATTCAGCCCTCTTCCCAACGCAATCTCAGCAC 780
Db 721 CACGGCCTGAGGATTGAGCCCA CCAATATTCAGCCCTCTTCCCAACGCAATCTCAGCAC 780
Qy 781 ATACGCGCGAGTACAACTAGCGCC CCAACCCGAGCAAAACACTGGATCATGCGCTACAGC 840
Db 781 ATACGCGCGAGTACAACTAGCGCC CCAACCCGAGCAAAACACTGGATCATGCGCTACAGC 840
Qy 841 GGGCCCATGAGCCCATTCACATGGA ATTCACCAACATGCTCAGCGGAGCGCTTGCAG 900
Db 841 GGGCCCATGAGCCCATTCACATGGA ATTCACCAACATGCTCAGCGGAGCGCTTGCAG 900
Qy 901 ACCCTCATGTGCGTGGAGCACTCCAT GAGACGATTTACAACTGTGGTTGAGACGGGC 960
Db 901 ACCCTCATGTGCGTGGAGCACTCCAT GAGACGATTTACAACTGTGGTTGAGACGGGC 960
Qy 961 GAGCTGACACACGATACATGCTATGAT TTTGACATCAGGGTCCGCTTCTACGTTG 1020
Db 961 GAGCTGACACACGATACATGCTATGAT TTTGACATCAGGGTCCGCTTCTACGTTG 1020
Qy 1021 GGCCTGTGAAAGGGAATCCATGCCATAT GATTTGACATCAGGGTCCGCTTCTACGTTG 1080
Db 1021 GGCCTGTGAAAGGGAATCCATGCCATAT GATTTGACATCAGGGTCCGCTTCTACGTTG 1080
Qy 1081 AGGGCCCCAACTGTGGAAGCGCGTGTCTG AATCCCAATCGTCTCAACATTTGACCTG 1140
Db 1081 AGGGCCCCAACTGTGGAAGCGCGTGTCTG AATCCCAATCGTCTCAACATTTGACCTG 1140
Qy 1141 GCGCCACCATCTCTGGACATTTGAGAGCTT GAGATACCTCGGATGAGCGGAAATCC 1200
Db 1141 GCGCCACCATCTCTGGACATTTGAGAGCTT GAGATACCTCGGATGAGCGGAAATCC 1200
Qy 1201 ATCTCAAGCTGTGGACACGGAGCGCGTGTG AATCGGTTTCACTTGAAGAAAGATG 1260
Db 1201 ATCTCAAGCTGTGGACACGGAGCGCGTGTG AATCGGTTTCACTTGAAGAAAGATG 1260
Qy 1261 AGGGCTGGGGGACCTCTTTGGTGGAGAGGCA GCTGCTACAAAGAGACAAAT 1320
Db 1261 AGGGCTGGGGGACCTCTTTGGTGGAGAGGCA GCTGCTACAAAGAGACAAAT 1320
Qy 1321 GACAAGTGGACGCCAGGAGGAGAACTTTCTC CCAAGTACAGAGTGTGAGGACCTG 1380
Db 1321 GACAAGTGGACGCCAGGAGGAGAACTTTCTC CCAAGTACAGAGTGTGAGGACCTG 1380
Qy 1381 TGTGAGCGTGTGATGATACAGACGGCGTGTGAG CAGCTGGGACAGAAAGTGGCAGTGTGTG 1440
Db 1381 TGTGAGCGTGTGATGATACAGACGGCGTGTGAG CAGCTGGGACAGAAAGTGGCAGTGTGTG 1440
Qy 1441 GAGGACGCCACGGGAGCTGGAAGTGTGATAGTG AAGTGAAGGGCCCATGCGGCTGGCGGC 1500
Db 1441 GAGGACGCCACGGGAGCTGGAAGTGTGATAGTG AAGTGAAGGGCCCATGCGGCTGGCGGC 1500
Qy 1501 AGCAGAGCCTCTTCCAACTCGTGGCCCAAGTA CTAAGGGCAGGGCAGCGGCTTGCACC 1560
Db 1501 AGCAGAGCCTCTTCCAACTCGTGGCCCAAGTA CTAAGGGCAGGGCAGCGGCTTGCACC 1560
Qy 1561 TGTGACAGCGGGACCTACAGCTCAGCCTGGCGG ACCTGGGCAAACTCTTCAAGAG 1620
Db 1561 TGTGACAGCGGGACCTACAGCTCAGCCTGGCGG ACCTGGGCAAACTCTTCAAGAG 1620
Qy 1621 AAGTCAAGGCCAGCTATGTCCGACGTCTCCATC CCGTCAAGTGGCCATCAGGTGGAC 1680
Db 1621 AAGTCAAGGCCAGCTATGTCCGACGTCTCCATC CCGTCAAGTGGCCATCAGGTGGAC 1680
Qy 1681 GGCAGGGTGTACACGCTAGGCTTGGTGTATGCGG CCCCAGAACTCTCAACAGCGG 1740
Db 1681 GGCAGGGTGTACACGCTAGGCTTGGTGTATGCGG CCCCAGAACTCTCAACAGCGG 1740

Qy 1741 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACA AGGATGTTGGTGGGACTTCAGTGGCACT 1800
Db 1741 CACTGGCCAGGGGCCCTTGAGGACCAAGATGACA AGGATGTTGGTGGGACTTCAGTGGCACT 1800
Qy 1801 GGAGGCTTCCGACTACTAGCGCCCAACCCCAAT TAAAGTGAACACATCGGTGTACATC 1860
Db 1801 GGAGGCTTCCGACTACTAGCGCCCAACCCCAAT TAAAGTGAACACATCGGTGTACATC 1860
Qy 1861 CTAGAGAACGACACAGTCCAGTCCAGTGTGACTG AACTGTACAGTCCCTGAGGCTTGGAAA 1920
Db 1861 CTAGAGAACGACACAGTCCAGTGTGACTGTGACTG AACTGTACAGTCCCTGAGGCTTGGAAA 1920
Qy 1921 GACCAACAGTGTGCAATGACCCAGCAAGATTGAAA CCGCTGCAGAACAAAAATTAAGAACCTG 1980
Db 1921 GACCAACAGTGTGCAATGACCCAGCAAGATTGAAA CCGCTGCAGAACAAAAATTAAGAACCTG 1980
Qy 1981 AGGGAAGTCCGAGTCCAGTCCAGTGTGACTGTG AAGTGAAGTGTGACTGTCAAAAATC 2040
Db 1981 AGGGAAGTCCGAGTCCAGTGTGACTGTGACTGTG AAGTGAAGTGTGACTGTCAAAAATC 2040
Qy 2041 AGCTACCAACCCAGCAAGGCGCTCAAGCACA GAGGCTCCAGTCTGCAATCTTTC 2100
Db 2041 AGCTACCAACCCAGCAAGGCGCTCAAGCACA GAGGCTCCAGTCTGCAATCTTTC 2100
Qy 2101 AGGAAGGCTTGCAGAGAGGACAGGTTGGCTGT TGGGAGCGAAGCGCAAGAG 2160
Db 2101 AGGAAGGCTTGCAGAGAGGACAGGTTGGCTGT TGGGAGCGAAGCGCAAGAG 2160
Qy 2161 AAACTCCGCAAGTGTCTCAAGCGCTCGAGAA CAAACGACACGTGCAGATGCCAGGCGCTC 2220
Db 2161 AAACTCCGCAAGTGTCTCAAGCGCTCGAGAA CAAACGACACGTGCAGATGCCAGGCGCTC 2220
Qy 2221 ACGTGCTTCAACCAAGCAACCGAGCACTGGCAG ACGCGCTTCTGGACACTGGGCGCT 2280
Db 2221 ACGTGCTTCAACCAAGCAACCGAGCACTGGCAG ACGCGCTTCTGGACACTGGGCGCT 2280
Qy 2281 TTCTGTGCTGCAACAGCGCAACATTAACACGTAC TGTGATGAGGACCAATGAG 2340
Db 2281 TTCTGTGCTGCAACAGCGCAACATTAACACGTAC TGTGATGAGGACCAATGAG 2340
Qy 2341 ACTCAAAATTTCTTCTGTGAAATTTGCAATTTG CAACTGGCTTCTAGAGTACTTGTATCTCAAC 2400
Db 2341 ACTCAAAATTTCTTCTGTGAAATTTGCAATTTG CAACTGGCTTCTAGAGTACTTGTATCTCAAC 2400
Qy 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACA CACTGAGACAGGATGCTCTCAACCCAG 2460
Db 2401 ACAGACCCCTACAGCTGATGATGAGTGAACACA CACTGAGACAGGATGCTCTCAACCCAG 2460
Qy 2461 CTACACGTACAGCTCATGAGCTGAGGAGCTGCAAG GGTTACAGAGTGAACCCCGG 2520
Db 2461 CTACACGTACAGCTCATGAGCTGAGGAGCTGCAAG GGTTACAGAGTGAACCCCGG 2520
Qy 2521 ACTCGAACAATGACCTGGGACTTAAAGATGGAAG AAGCTATGAGCAATACAGGCACTTT 2580
Db 2521 ACTCGAACAATGACCTGGGACTTAAAGATGGAAG AAGCTATGAGCAATACAGGCACTTT 2580
Qy 2581 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCT TCCAAATCACTGGGCAAACTGTGG 2640
Db 2581 CAGGCTCGAAAGTGGCCAGAAATGAAGAGACCTTTCT TCCAAATCACTGGGCAAACTGTGG 2640
Qy 2641 GAAGGCTGGGAAGGTTAAGAAAACAAGAGGTGGA CTTCCAAAAATAGAGGCAATCAAC 2700
Db 2641 GAAGGCTGGGAAGGTTAAGAAAACAAGAGGTGGA CTTCCAAAAATAGAGGCAATCAAC 2700
Qy 2701 TGACTGCACAGGCAATGAAGAAACCATGTGGTGAAT TCCAGGACAGCTGTGCTATTGGCC 2760
Db 2701 TGACTGCACAGGCAATGAAGAAACCATGTGGTGAAT TCCAGGACAGCTGTGCTATTGGCC 2760
Qy 2761 AGGAGGCTTGAAGAAAGCAAGCAGCACTCTCAGTCA AATGACAGATTTCTGGAGGATAAC 2820
Db 2761 AGGAGGCTTGAAGAAAGCAAGCAGCACTCTCAGTCA AATGACAGATTTCTGGAGGATAAC 2820

Db 3417 GTGACTATTTCATGTAATCAACATGGGAACCTTTTAGGGGAACCTAATAAGAAATCCCAA 3476
QY 3301 TTTTCAGGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3352
Db 3477 TTTTCAGGAGTGGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3528
RESULT 4
US-09-970-287-11
; Sequence 11, Application US/09970287
; Publication No. US20030166283A1
; GENERAL INFORMATION:
; APPLICANT: GLUCKSMANN, Maria A.
; APPLICANT: MEYERS, Rachel
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; APPLICANT: SILOS-SANTIAGO, Inmaculada
; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR
; FILE REFERENCE: 10147-61U1
; CURRENT APPLICATION NUMBER: US/09/970,287
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US 60/257,082
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 4397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-970-287-11
Query Match 99.5%; Score 3347.2; DB 13; Length 4397;
Best Local Similarity 99.9%; Pred No. 0;
Matches 3349; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GACTCCCGATCCCAAGAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 60
Db 286 GACTCCCGATCCCAAGAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 345
QY 61 CTCGTGCTGTCTGCTCGCACTGTGTCTCCCTGCTGGTGGAGTGGCCCTTC 120
Db 346 CTCGTGCTGTCTGCTCGCACTGTGTCTCCCTGCTGGTGGAGTGGCCCTTC 405
QY 121 CTGTCCGACCAACCGCTTGAAGGAGGAGTTTCAGAGGAGCCGACAGAAACATCCGCCCAAC 180
Db 406 CTGTCCGACCAACCGCTTGAAGGAGGAGTTTCAGAGGAGCCGACAGAAACATCCGCCCAAC 465
QY 181 ATCATCTGTGTCTGACGAGAGACAGAGATGTGAGTGGTTCATGCGAGTGAAC 240
Db 466 ATCATCTGTGTCTGACGAGAGACAGAGATGTGAGTGGTTCATGCGAGTGAAC 525
QY 241 AAGACCCGGCGCATCATGGAGCAGGGCGGGCGCACTTCATCAACGCTTCGTGACCA 300
Db 526 AAGACCCGGCGCATCATGGAGCAGGGCGGGCGCACTTCATCAACGCTTCGTGACCA 585
QY 301 CCATGTGTCCTCAGCTTCCTCATCTCATCTGCAAGTACGTCCCAACCAAC 360
Db 586 CCATGTGTCCTCAGCTTCCTCATCTCATCTGCAAGTACGTCCCAACCAAC 645
QY 361 ACCTACCAACATGAACTGCTCTCGCCCTCTGCGAGGACAGCAGCAGAGCCG 420
Db 646 ACCTACCAACATGAACTGCTCTCGCCCTCTGCGAGGACAGCAGCAGAGCCG 705
QY 421 ACCTTTCGGTGTACCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTT 480
Db 706 ACCTTTCGGTGTACCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCTT 765
QY 481 AATGATACACGGCTCTAGTGCACCGCTCTGGAAGAGTGGTGCAGCTCTTAAA 540
Db 766 AATGATACACGGCTCTAGTGCACCGCTCTGGAAGAGTGGTGCAGCTCTTAAA 825
QY 541 AACTCCCGCTTTATATACTACACGCTGTGTGCGAAGCGGGTGAAGAAAGACGCGCTCC 600

Db 826 AACTCCCGCTTTTATAACTACACGCTGTGTGCGAAGCGGGTGAAGAGAACGACGCTCC 885
QY 601 GACTACTCCAGGATTTACCTCAGAGACCTCATCAACCAATGACAGGGTGAAGCTTCTTCGG 660
Db 886 GACTACTCCAGGATTTACCTCAGAGACCTCATCAACCAATGACAGGGTGAAGCTTCTTCGG 945
QY 661 ACGTCCAAGAAGATGTACCCGACAGGCGCAGTCTCTCATGCTCATCAGCCATCAGAGCCGCC 720
Db 946 ACGTCCAAGAAGATGTACCCGACAGGCGCAGTCTCTCATGCTCATCAGCCATCAGAGCCGCC 1005
QY 721 CAGGCGCTCAGGATTTACCCGACCAATATTACGGCTCTTCCCAACGCAATCTCAGAC 780
Db 1006 CAGGCGCTCAGGATTTACCCGACCAATATTACGGCTCTTCCCAACGCAATCTCAGAC 1065
QY 781 ATCAGCGCCGAGCTACAACTACGCGCCCAACCCGCAACAACTGGATCATGCGCTACAG 840
Db 1066 ATCAGCGCCGAGCTACAACTACGCGCCCAACCCGCAACAACTGGATCATGCGCTACAG 1125
QY 841 GGGCCCATGAGCCCATCCATGGAATTCACAAATGCTCCAGCGGAAGCGCTTGCA 900
Db 1126 GGGCCCATGAGCCCATCCATGGAATTCACAAATGCTCCAGCGGAAGCGCTTGCA 1185
QY 901 ACCCTCATGTCGTGGAGACTCCATGGAGAGATTTCACACATGCTGTTGAGACGGGC 960
Db 1186 ACCCTCATGTCGTGGAGACTCCATGGAGAGATTTCACACATGCTGTTGAGACGGGC 1245
QY 961 GAGCTGACAAACGCTATCGTATACACCGCCGACACCGGTTACCAATCGGCCAGTTT 1020
Db 1246 GAGCTGACAAACGCTATCGTATACACCGCCGACACCGGTTACCAATCGGCCAGTTT 1305
QY 1021 GGCCTGTTGAAAGGAAATCCATGCCATATGATTTGACATCAGGGTCCCGTTCTACGTT 1080
Db 1306 GGCCTGTTGAAAGGAAATCCATGCCATATGATTTGACATCAGGGTCCCGTTCTACGTT 1365
QY 1081 AGGGGCCCCAAGCTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCTCAACATTCACCTG 1140
Db 1366 AGGGGCCCCAAGCTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCTCAACATTCACCTG 1425
QY 1141 GCGCCCAACCATCTGGAATTCAGAGGCTTGACATACCTGCGGATATGACGCGGAATTC 1200
Db 1426 GCGCCCAACCATCTGGAATTCAGAGGCTTGACATACCTGCGGATATGACGCGGAATTC 1485
QY 1201 ATCTCTCAAGCTCTGGAACACGAGCGGCGGTTGAACTCGTTTCACTTGAAAAGAGATG 1260
Db 1486 ATCTCTCAAGCTCTGGAACACGAGCGGCGGTTGAACTCGTTTCACTTGAAAAGAGATG 1545
QY 1261 AGGCTCTGCGGGAATCTCTTGTGTGAGAGAGCAAGTCTGTCTACAAAGAGACAAT 1320
Db 1546 AGGCTCTGCGGGAATCTCTTGTGTGAGAGAGCAAGTCTGTCTACAAAGAGACAAT 1605
QY 1321 GACAGGTGAGCGCCGAGGAGAGAACTTCTGCGCAAGTACCAAGTGTGAGAGACCTG 1380
Db 1606 GACAGGTGAGCGCCGAGGAGAGAACTTCTGCGCAAGTACCAAGTGTGAGAGACCTG 1665
QY 1381 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGAGCTGGGACAGAACTGGCAGTGTGTG 1440
Db 1666 TGTGAGCGTGTGAGTACCAAGCGGCTGTGAGAGCTGGGACAGAACTGGCAGTGTGTG 1725
QY 1441 GAGGAGCGCAACGGGAGAGTGAAGTGCATAGTGAAGGGGCCCATCGGCTGGGCGGC 1500
Db 1726 GAGGAGCGCAACGGGAGAGTGAAGTGCATAGTGAAGGGGCCCATCGGCTGGGCGGC 1785
QY 1501 AGCAGAGCGCTCTCCAACTCTGCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGACCC 1560
Db 1786 AGCAGAGCGCTCTCCAACTCTGCGCCCAAGTACTACGGGCGAGGCGAGGCGCTGACCC 1845
QY 1561 TGTGACAGCGGGAATCAAGCTAGCTGCGCGGACCGCGGAAAAAACTCTTCAAGAAG 1620
Db 1846 TGTGACAGCGGGAATCAAGCTAGCTGCGCGGACCGCGGAAAAAACTCTTCAAGAAG 1905
QY 1621 AAGTACAAGGCCAGCTATGTCCGAGTCGCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1680
Db 1906 AAGTACAAGGCCAGCTATGTCCGAGTCGCTCCATCCGCTCAGTGGCCATCGAGTGGAC 1965

1681 GGCAGGGTGATACAGTAGGCTGGGTGATGCGCCAGCGCCCGAAGCCTCAACAGCGG 1740
1666 GGCAGGGTGATACAGTAGGCTGGGTGATGCGCCAGCGCCCGAAGCCTCAACAGCGG 2025
1741 CACTGGCAGGGGCGCCCTGAGGACCAAGATGACAGAGATGAGGAGCTTCAGTGGCACT 1800
2026 CACTGGCAGGGGCGCCCTGAGGACCAAGATGACAGAGATGAGGAGCTTCAGTGGCACT 2085
1801 GGAGGCTTCCGACTACTACGCGCCCAACCCATTAAGTGAACATCGGTGTACATC 1860
2086 GGAGGCTTCCGACTACTACGCGCCCAACCCATTAAGTGAACATCGGTGTACATC 2145
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTGAGGCGCTGGA 1920
2146 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTGAGGCGCTGGA 2205
1921 GACCAACAGCTGCATCGACACAGAGATGAACCCCTGCAGACAAATTAAGAACCTG 1980
2206 GACCAACAGCTGCATCGACACAGAGATGAACCCCTGCAGACAAATTAAGAACCTG 2265
1981 AGGGAAGTCCGAGGTCACTGAGGAAAGGCGGCCAGAGAGATGTGACTGTCAAAAATC 2040
2266 AGGGAAGTCCGAGGTCACTGAGGAAAGGCGGCCAGAGAGATGTGACTGTCAAAAATC 2325
2041 AGCTACCAACCCAGGACAAAGGCGCCCTCAAGCAGAGAGTCCAGTGTGCTCCTTTC 2100
2326 AGCTACCAACCCAGGACAAAGGCGCCCTCAAGCAGAGAGTCCAGTGTGCTCCTTTC 2385
2101 AGGAAGGGCTCCAGAGAGAGGACAAAGTGTGGCTGTGCGGAGCAGAGCGAAGAG 2160
2386 AGGAAGGGCTCCAGAGAGAGGACAAAGTGTGGCTGTGCGGAGCAGAGCGAAGAG 2445
2161 AAATCCGCAAGCTGCTAAAGCGCTGAGAACAAAGCAGCAGTGCAGATGCCAGGCTC 2220
2446 AAATCCGCAAGCTGCTAAAGCGCTGAGAACAAAGCAGCAGTGCAGATGCCAGGCTC 2505
2221 AGTGCTTCCACCCAGACAAACAGCAGTGGCAGAGCGCGCTTTCGACACTGGGCGCT 2280
2506 AGTGCTTCCACCCAGACAAACAGCAGTGGCAGAGCGCGCTTTCGACACTGGGCGCT 2565
2281 TTCTGTGCTGCACAGCGCCCAACAAATACAGTACTGTGTCATGAGGACCAATCAATGAG 2340
2566 TTCTGTGCTGCACAGCGCCCAACAAATACAGTACTGTGTCATGAGGACCAATCAATGAG 2625
2341 ACTCAAAATTTCTCTCTGTAATTTGCAATGCTTCTTCTAGAGTACTTTGATCTCAAC 2400
2626 ACTCAAAATTTCTCTCTGTAATTTGCAATGCTTCTTCTAGAGTACTTTGATCTCAAC 2685
2401 ACAGACCCCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
2686 ACAGACCCCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2745
2461 CTACAGCTACAGCTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2520
2746 CTACAGCTACAGCTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2805
2521 ACTCGAAACATGCACTGGGCTTAAAGATGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2580
2806 ACTCGAAACATGCACTGGGCTTAAAGATGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2865
2581 CAGCGTGAAGTGGCAGAAATGAAGAGACTTTTCCAAATCACTGGGCAACTGTGG 2640
2866 CAGCGTGAAGTGGCAGAAATGAAGAGACTTTTCCAAATCACTGGGCAACTGTGG 2925
2641 GAAGGCTGGAGAGTTTAAAGAAACAGAGGTTGGACCTCCAAAACATAGAGGATCAAC 2700
2926 GAAGGCTGGAGAGTTTAAAGAAACAGAGGTTGGACCTCCAAAACATAGAGGATCAAC 2985
2701 TGACTGCACAGGCAATGAAAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
2986 TGACTGCACAGGCAATGAAAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3045

2761 AGGAGGCTTGAGAAAGCAGCAGCAGCTCTCAGTCAATGACAGATCTTGTGAGATAAC 2820
3046 AGGAGGCTTGAGAAAGCAGCAGCAGCTCTCAGTCAATGACAGATCTTGTGAGATAAC 3105
2821 CAGCAGGACGAGATTAATCTCAGGAAGTCCATTTTGTGCCCCCTGCTTTGCTTTGATTA 2880
3106 CAGCAGGACGAGATTAATCTCAGGAAGTCCATTTTGTGCCCCCTGCTTTGCTTTGATTA 3165
2881 TACCTCAGCAGCTGCACAAATGCAATTTTGTATCAAAAAGTCAACCACTAACCTCC 2940
3166 TACCTCAGCAGCTGCACAAATGCAATTTTGTATCAAAAAGTCAACCACTAACCTCC 3225
2941 CCAGAAGCTCAAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
3226 CCAGAAGCTCAAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3285
3001 CCCAAGGGGAAAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAA 3060
3286 CCCAAGGGGAAAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAA 3345
3061 TCCTCTTATTCTTTTGGTTTGTCAAAAAGAGAACTAAGAGCAGGACAGAGGCAACGT 3120
3346 TCCTCTTATTCTTTTGGTTTGTCAAAAAGAGAACTAAGAGCAGGACAGAGGCAACGT 3405
3121 GAGAGGCTGAAACAGTCAGAGAGCTTTGACATGAGTCAATGAGCAGACAAAGAGATGA 3180
3406 GAGAGGCTGAAACAGTCAGAGAGCTTTGACATGAGTCAATGAGCAGACAAAGAGATGA 3465
3181 CATTTACCTAGCAGTATAAACCTCTGCTCTCTGAGAACTGCTTCATTGTATATAT 3240
3466 CATTTACCTAGCAGTATAAACCTCTGCTCTCTGAGAACTGCTTCATTGTATATAT 3525
3241 GTGACTATTACATGATTAATCAACATGAGGAACTTTTAGGGAACCTAATAAGAAATCCCAA 3300
3526 GTGACTATTACATGATTAATCAACATGAGGAACTTTTAGGGAACCTAATAAGAAATCCCAA 3585
3301 TTTTTCAGGAGTGGTGTGTCATTAATTAACGCTCTGTGCGCAGTGTAAAAAGAAA 3352
3586 TTTTTCAGGAGTGGTGTGTCATTAATTAACGCTCTGTGCGCAGTGTAAAAAGAAA 3637

RESULT 5
US-10-025-966A-13
; Sequence 13, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; TITLE OF INVENTION: Sulfatases and methods of use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025,966A
; PRIOR FILING DATE: 2002-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 4286
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-966A-13

Query Match 99.5%; Score 3346.8; DB 13; Length 4286;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
1 GACTCCCGCATCCAAAAGAACCAACAGATCAGCAAAAAGAAAGATGGCCCCCGAGC 60
632 GACTCCCGCATCCAAAAGAACCAACAGATCAGCAAAAAGAAAGATGGCCCCCGAGC 691

QY 541 AACTCCGCTTTTATACCTACACGCTGTGTCGGAACGGGGTGAAGAAAGACAGCGCTCC 600
DB 1172 AACTCCGCTTTTATACCTACACGCTGTGTCGGAACGGGGTGAAGAAAGACAGCGCTCC 1231
QY 601 GACTACTCCAAAGATTACTTCACAGACTTCATCAACCAATGACAGCTGAGCTTCTCCGC 660
DB 1232 GACTACTCCAAAGATTACTTCACAGACTTCATCAACCAATGACAGCTGAGCTTCTCCGC 1291
QY 661 AGCTCAAGAAAGATTACCCGACAGCCGACAGGCGAGTCTCTATGGTTCATCAGCCATCGAGCCCC 720
DB 1292 AGCTCAAGAAAGATTACCCGACAGCCGACAGGCGAGTCTCTATGGTTCATCAGCCATCGAGCCCC 1351
QY 721 CACGCGCTGAGGATTCAGCCCCCAATATTCAGCCCTCTCCAAAGCAATCTCAGAC 780
DB 1352 CACGCGCTGAGGATTCAGCCCCCAATATTCAGCCCTCTCTCCAAAGCAATCTCAGAC 1411
QY 781 ATCAGCCGAGCTACAACCTACGCGCCCAACCGGACAAACACTGATCATCGCTACACG 840
DB 1412 ATCAGCCGAGCTACAACCTACGCGCCCAACCGGACAAACACTGATCATCGCTACACG 1471
QY 841 GGGCCCATGAGCCCATCGACATGGAATTACCAACATGCTCCAGCGGAAGCGTTGAG 900
DB 1472 GGGCCCATGAGCCCATCGACATGGAATTACCAACATGCTCCAGCGGAAGCGTTGAG 1531
QY 901 ACCCTCATGTCGGTGGAGCACTCCATGAGACGATTTACCAACATGCTGTTGACAGCGG 960
DB 1532 ACCCTCATGTCGGTGGAGCACTCCATGAGACGATTTACCAACATGCTGTTGAGAGCGG 1591
QY 961 GAGCTGGACACACGTATACATCGTATACACCGCGACACGTTTACCAATCGGCGAGTTT 1020
DB 1592 GAGCTGGACACACGTATACATCGTATACACCGCGACACGTTTACCAATCGGCGAGTTT 1651
QY 1021 GGCCTGGTGAAGGGAATCCATGCTATATGAGTTGACATCAGGCTCCGCTTACGTTG 1080
DB 1652 GGCCTGGTGAAGGGAATCCATGCTATATGAGTTGACATCAGGCTCCGCTTACGTTG 1711
QY 1081 AGGGCCCCAAGCTGGAAAGCGGCTGTGTAATCCCACTCGTCTCAACATGACCTG 1140
DB 1712 AGGGCCCCAAGCTGGAAAGCGGCTGTGTAATCCCACTCGTCTCAACATGACCTG 1771
QY 1141 GCGCCCATGCTGGACATTCAGGCGCTGACATACCTGCGGATGAGAGGGAATCC 1200
DB 1772 GCGCCCATGCTGGACATTCAGGCGCTGACATACCTGCGGATGAGAGGGAATCC 1831
QY 1201 ATCTCTAAGCTGCTGGACACAGGCGCGGCTGTAATCGTTTCACTTGAAGAAAGAGATG 1260
DB 1832 ATCTCTAAGCTGCTGGACACAGGCGCGGCTGTAATCGTTTCACTTGAAGAAAGAGATG 1891
QY 1261 AGGCTCTGGCGGACTCCTTCTTGGTGAGAGAGCAAGCTGCTTACACAGAGAGCAAT 1320
DB 1892 AGGCTCTGGCGGACTCCTTCTTGGTGAGAGAGCAAGCTGCTTACACAGAGAGCAAT 1951
QY 1321 GACAGGTGGAGCGCCAGGAGGAGCACTTCTGCGCAAGTACAGCGTGTGAGGACCTG 1380
DB 1952 GACAGGTGGAGCGCCAGGAGGAGCACTTCTGCGCAAGTACAGCGTGTGAGGACCTG 2011
QY 1381 TGTGAGCGTGTGAGTACAGACGCGGCTGTGAGAGCTGGGACAGAGTGGCAGTGTG 1440
DB 2012 TGTGAGCGTGTGAGTACAGACGCGGCTGTGAGAGCTGGGACAGAGTGGCAGTGTG 2071
QY 1441 GAGAGCGCCAGGCGGAGCTGAAGCTGATAGTGAAGTGAAGGCCCTTCGCGCTGGCGGC 1500
DB 2072 GAGAGCGCCAGGCGGAGCTGAAGCTGATAGTGAAGTGAAGGCCCTTCGCGCTGGCGGC 2131
QY 1501 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTACGGGAGGCGAGCGAGGCTGCACC 1560
DB 2132 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTACGGGAGGCGAGCGAGGCTGCACC 2191
QY 1561 TGTGACAGCGGGAATCAAGCTCAGCTGCGGACGCGGGAAGAACTCTTCAAGAG 1620
DB 2192 TGTGACAGCGGGAATCAAGCTCAGCTGCGGACGCGGGAAGAACTCTTCAAGAG 2251
QY 1621 AAGTACAGGCGAGCTATGTCGCGAGTCTCCATCCGCTCAGTGGGCCATCGAGGTGGAC 1680

DB 2252 AAGTACAGGCGAGCTATGTCGCGAGTCTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2311
QY 1681 GGCAGGGTGTACCACTGAGGCTGGTGTATGCGGCCACCGCCGAAACCTCACAAGCGG 1740
DB 2312 GGCAGGGTGTACCACTGAGGCTGGTGTATGCGGCCACCGCCGAAACCTCACAAGCGG 2371
QY 1741 CACTGGCCCAAGGGGCCCTCAGGACCAAGATGACAAGATGGTGGGACTTCAGTGGCACT 1800
DB 2372 CACTGGCCCAAGGGGCCCTCAGGACCAAGATGACAAGATGGTGGGACTTCAGTGGCACT 2431
QY 1801 GAGGGCTTCCGACTACTCAGCGCAACCCCATTTAAAGTGACACATCGGTGTACATC 1860
DB 2432 GAGGGCTTCCGACTACTCAGCGCAACCCCATTTAAAGTGACACATCGGTGTACATC 2491
QY 1861 CTAGAGAAACGACACAGTGTGAGTGTGACCTGGACCTGTACAAGTCTCGAGGCTGGAAA 1920
DB 2492 CTAGAGAAACGACACAGTGTGAGTGTGACCTGGACCTGTACAAGTCTCGAGGCTGGAAA 2551
QY 1921 GACCAACAGCTGCATCATCGACCAAGATTTGAAACCTTCAGAACAAATTAAGAACCTG 1980
DB 2552 GACCAACAGCTGCATCATCGACCAAGATTTGAAACCTTCAGAACAAATTAAGAACCTG 2611
QY 1981 AGGGAAGTCCGAGGTCACTGGAAGAAAGCGCCAGAGATGTGACTGTCACAAAATC 2040
DB 2612 AGGGAAGTCCGAGGTCACTGGAAGAAAGCGCCAGAGATGTGACTGTCACAAAATC 2671
QY 2041 AGCTACCAACCCAGCACAAAGCCGCTCAAGCACAGAGGCTCCAGTCTGCTCTTTC 2100
DB 2672 AGCTACCAACCCAGCACAAAGCCGCTCAAGCACAGAGGCTCCAGTCTGCTCTTTC 2731
QY 2101 AGGAGGGCTCCAGAGNAGAGCAAGGTGTGCTGTTGCGGAGCAGAGCGCAAGAG 2160
DB 2732 AGGAGGGCTCCAGAGNAGAGCAAGGTGTGCTGTTGCGGAGCAGAGCGCAAGAG 2791
QY 2161 AAATCCGCAAGCTGTCAAGCGCTGAGAACACAGACACGTGAGCATGCGAGGCTC 2220
DB 2792 AAATCCGCAAGCTGTCAAGCGCTGAGAACACAGACACGTGAGCATGCGAGGCTC 2851
QY 2221 AGTGTCTTCCACCAAGCAACAGCACTGGCAGACGGCGCTTCTGCACTTGGGGCT 2280
DB 2852 AGTGTCTTCCACCAAGCAACAGCACTGGCAGACGGCGCTTCTGCACTTGGGGCT 2911
QY 2281 TTCTGTGCTGACAGCGCCCAACCAATTAACGTACTGTCATGAGGACCAATCAATGAG 2340
DB 2912 TTCTGTGCTGACAGCGCCCAACCAATTAACGTACTGTCATGAGGACCAATCAATGAG 2971
QY 2341 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2400
DB 2972 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 3031
QY 2401 ACAGACCCCTACAGCTGATGATGACAGTGAACACACTGGACAGGATGTCTCAACCAG 2460
DB 3032 ACAGACCCCTACAGCTGATGATGACAGTGAACACACTGGACAGGATGTCTCAACCAG 3091
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 2520
DB 3092 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGG 3151
QY 2521 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 2580
DB 3152 ACTCAAAATTTCTTCTGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC 3211
QY 2581 CAGCGTCAAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640
DB 3212 CAGCGTCAAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 3271
QY 2641 GAAGCTGGGAGGTTAAGAAACACAGAGGTGGACCTTCCAAAAACATAGAGGATCAACC 2700
DB 3272 GAAGCTGGGAGGTTAAGAAACACAGAGGTGGACCTTCCAAAAACATAGAGGATCAACC 3331
QY 2701 TGACTGCACAGGCAATGAAGAAACCAATGAGGAGTGTGCTGCTTATTTGGCC 2760


```
Db 3332 TGACGACAGGCAATGAATAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3391
Qy AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3451
Qy 2821 CAGCAGAGCAGAGATTAATCTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880
Db 3452 CAGCAGAGCAGAGATTAATCTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3511
Qy 2881 TACCTCACAGCTGCACAAATGCAATTTTCTGATCAAAAAGTCAACCTAACCCCTCCC 2940
Db 3512 TACCTCACAGCTGCACAAATGCAATTTTCTGATCAAAAAGTCAACCTAACCCCTCCC 3571
Qy 2941 CCAGAACTCACAAGGAAACCGAGAGAGCGAGCGAGAGATTTCTTGGAAATTTCT 3000
Db 3572 CCAGAACTCACAAGGAAACCGAGAGAGCGAGCGAGAGATTTCTTGGAAATTTCT 3631
Qy 3001 CCCAAGGCGGAAGTCAATTCGAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAA 3060
Db 3632 CCCAAGGCGGAAGTCAATTCGAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAA 3691
Qy 3061 TCCTCTTATTTCTTTGGTTTGTACAAAGAAAGAACTAAGAACGAGCAGAGGCAACGT 3120
Db 3692 TCCTCTTATTTCTTTGGTTTGTACAAAGAAAGAACTAAGAACGAGCAGAGGCAACGT 3751
Qy 3121 GGAGAGGCTGAAACAGTGCAGAGAGCTTTGACAAATGAGTCACTAGCAGCAGAAAGAGATGA 3180
Db 3752 GGAGAGGCTGAAACAGTGCAGAGAGCTTTGACAAATGAGTCACTAGCAGCAGAAAGAGATGA 3811
Qy 3181 CATTTACCTAGCACTATAACCCCTGGTTGCTCTGAAGAACTGCTTCAATTTGATATAT 3240
Db 3812 CATTTACCTAGCA-TATAACCCCTGGTTGCTCTGAAGAACTGCTTCAATTTGATATAT 3870
Qy 3241 GTGACTATTTTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATGAAGAAATCCCAA 3300
Db 3871 GTGACTATTTTACATGTAATCAACATGGGAACCTTTAGGGGAACCTTAATGAAGAAATCCCAA 3930
Qy 3301 TTTTCAGGAGTGTGTGTCTCAATAAACGCTCTGTGGCCAGTGTAAAGGAAAGAAAGAAA 3360
Db 3931 TTTTCAGGAGTGTGTGTCTCAATAAACGCTCTGTGGCCAGTGTAAAGGAAAGAAAGAAA 3990
Qy 3361 AA 3362
Db 3991 AA 3992

RESULT 7
; Sequence 17, Application US/10182951
; Publication No. US20030138895A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SHIH, Leo L.
; APPLICANT: YANG, Junming
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0033 PCT
; CURRENT APPLICATION NUMBER: US/10/182,951
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/181,856; 10/183,684; 60/185,141; 60/186,818; 60/188,345;
; PRIOR FILING DATE: 2000-02-11; 2000-02-17; 2000-02-25; 2000-03-03; 2000-03-09; 2000-0
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 17
```

```
; LENGTH: 3348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030138895A1 1558210CB1
US-10-182-951-17

Query Match          99.5%; Score 3344.8; DB 13; Length 3348;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 12 CCCAAAGAACACACAGATCAGCAAAAAGAGATGGCCCCCGAGCCTCGTGTGTG 71
Db 1 CCCAAAGAACACACAGATCAGCAAAAAGAGATGGCCCCCGAGCCTCGTGTGTG 60
Qy 72 CTTGCTGTCGCAACTGTCTCTCCCTGCTGGGTGGAAGCTCGGCTTCTGTGCAACA 131
Db 61 CTTGCTGTCGCAACTGTCTCTCCCTGCTGGGTGGAAGCTCGGCTTCTGTGCAACA 120
Qy 132 CCGCTCGAAAAGGCGAGGTTTCAGAGGGACCGCAGAACATCCGCCCCCAACATCATCTCGT 191
Db 121 CCGCTCGAAAAGGCGAGGTTTCAGAGGGACCGCAGAACATCCGCCCCCAACATCATCTCGT 180
Qy 192 GCTCAACGAGCAGACAGATGTGAGCTGGTTCATGCAAGTGTATGAACAAAGACCCGGCG 251
Db 181 GCTCAACGAGCAGACAGATGTGAGCTGGTTCATGCAAGTGTATGAACAAAGACCCGGCG 240
Qy 252 CATCATGAGCAGAGGGCGGCGCACTTTCATCAACGCTTCGTGACCAACACCCATGTGCTG 311
Db 241 CATCATGAGCAGAGGGCGGCGCACTTTCATCAACGCTTCGTGACCAACACCCATGTGCTG 300
Qy 312 CCCCTCAGCTCTCTCATCTCTCACTGGCAAGTACGTCACCAACCAACCAACCTTACACCAA 371
Db 301 CCCCTCAGCTCTCTCATCTCTCACTGGCAAGTACGTCACCAACCAACCAACCTTACACCAA 360
Qy 372 CAATGAGAACTGCTCTCTCGCCCTCTCGGAGGACACAGCAGAGCGGCGACCTTTCGCGT 431
Db 361 CAATGAGAACTGCTCTCTCGCCCTCTCGGAGGACACAGCAGAGCGGCGACCTTTCGCGT 420
Qy 432 GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTTCGGGAAGTATCTTAATGAATACAA 491
Db 421 GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTTCGGGAAGTATCTTAATGAATACAA 480
Qy 492 CGGCTCCTAGTCGCCACCGGCTGGAGGAGTGGTTCGCACTCTTAAAAAATCTCCGCTT 551
Db 481 CGGCTCCTAGTCGCCACCGGCTGGAGGAGTGGTTCGCACTCTTAAAAAATCTCCGCTT 540
Qy 552 TTATACTACACGCTGTGTGCGAAACGGGGTGAAGAAAGAACAGCAGGCTCCGACTACTCCAA 611
Db 541 TTATACTACACGCTGTGTGCGAAACGGGGTGAAGAAAGAACAGCAGGCTCCGACTACTCCAA 600
Qy 612 GGATTACCTCAGACAGCTCATACCAATGACAGCTGAGCTTCTCCGACGTCCTCAAGAA 671
Db 601 GGATTACCTCAGACAGCTCATACCAATGACAGCTGAGCTTCTCCGACGTCCTCAAGAA 660
Qy 672 GATGTACCCGACAGGCGCAGTCTCTCATGTCTATGCTATGAGTATGAGTATGAGTATGAGT 731
Db 661 GATGTACCCGACAGGCGCAGTCTCTCATGTCTATGCTATGAGTATGAGTATGAGTATGAGT 720
Qy 732 GATGTACCCGACAGGCGCAGTCTCTCATGTCTATGCTATGAGTATGAGTATGAGTATGAGT 791
Db 721 GATGTACCCGACAGGCGCAGTCTCTCATGTCTATGAGTATGAGTATGAGTATGAGTATGAGT 780
Qy 792 CTACAACTACGCGCCCAACCCGAGCAAAACACTGGATCATGCGCTACACGGGGCCCATGAA 851
Db 781 CTACAACTACGCGCCCAACCCGAGCAAAACACTGGATCATGCGCTACACGGGGCCCATGAA 840
Qy 852 GCCCATCCCATGGAATTCACCAACATGTCTCCAGCGGAAGCGCTTTCAGACCCCTCATGTC 911
Db 841 GCCCATCCCATGGAATTCACCAACATGTCTCCAGCGGAAGCGCTTTCAGACCCCTCATGTC 900
Qy 912 GGTGGACGACTCCATGAGAGCGATTTTACAACTATGCTGTTGAGACGGGCGAGCTGACAA 971
```

Db 901 GGTGGACGACTCCATGGAGCAGATTTCACCAATGCTGGTGGAGCGGCGAGCTGGACAA 960
QY 972 CAGCTACATCGTATACACCGCGGACCAACCGTTTACACATCGGCGGAGTTTGGGCTCGTGA 1031
Db 961 CAGCTACATCGTATACACCGCGGACCAACCGTTTACACATCGGCGGAGTTTGGGCTCGTGA 1020
QY 1032 AGGGAATCCATGCCATATAGTTTGACATCAGGGTCCGTTCTACTAGTGGGGCCCCAA 1091
Db 1021 AGGGAATCCATGCCATATAGTTTGACATCAGGGTCCGTTCTACTAGTGGGGCCCCAA 1080
QY 1092 CGTGAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTGACCTGGGCCCCACCAT 1151
Db 1081 CGTGAAGCGGCTGTCTGAATCCCAATCGTCTCAACATTGACCTGGGCCCCACCAT 1140
QY 1152 CTTGGACATTCAGGCTCGGATCATCTCGGATATGACCGGGAATCATCTCTCAAGCT 1211
Db 1141 CTTGGACATTCAGGCTCGGATCATCTCGGATATGACCGGGAATCATCTCTCAAGCT 1200
QY 1212 GCTGGACCGGAGCGGCTGAATCGGTTTCACTTGAAGAAGAGATGAGGTTGCGG 1271
Db 1201 GCTGGACCGGAGCGGCTGAATCGGTTTCACTTGAAGAAGAGATGAGGTTGCGG 1260
QY 1272 GGAATCCTTCTTGTGGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGA 1331
Db 1261 GGAATCCTTCTTGTGGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGA 1320
QY 1332 CGCCAGAGAGAGAACTTCTGCCCAAGTACACGCGTGTGAAGGACCTGTGTGACGCTGC 1391
Db 1321 CGCCAGAGAGAGAACTTCTGCCCAAGTACACGCGTGTGAAGGACCTGTGTGACGCTGC 1380
QY 1392 TGAGTACACAGCGGCTGTGACAGCTGGGACAGAAAGTGGCAGTGTGTGGAGGAGCGCAC 1451
Db 1381 TGAGTACACAGCGGCTGTGACAGCTGGGACAGAAAGTGGCAGTGTGTGGAGGAGCGCAC 1440
QY 1452 GGGGAAGCTGAAGCTGCATTAAGTGAAGGCGGCGGCGGCTGGGCGGCGACGAGCGCT 1511
Db 1441 GGGGAAGCTGAAGCTGCATTAAGTGAAGGCGGCGGCGGCTGGGCGGCGACGAGCGCT 1500
QY 1512 CTCCAACTCGTGGCCCAAGTACTACGGGCGGCGGAGCGGCGGCTGGGCGGCGGCGG 1571
Db 1501 CTCCAACTCGTGGCCCAAGTACTACGGGCGGCGGAGCGGCGGCTGGGCGGCGGCGG 1560
QY 1572 GGACTACAAGCTCAGCTGGCGGCGGCGGAGAACTTTCAGGAAGAGTACAAGGC 1631
Db 1561 GGACTACAAGCTCAGCTGGCGGCGGCGGAGAACTTTCAGGAAGAGTACAAGGC 1620
QY 1632 CAGCTATGTCGAGCTCGTCCATCGCTCAGTGGCCATCAGGTGACGCGGAGGTTGA 1691
Db 1621 CAGCTATGTCGAGCTCGTCCATCGCTCAGTGGCCATCAGGTGACGCGGAGGTTGA 1680
QY 1692 CCAGTATGTCGAGCTCGTCCATCGCTCAGTGGCCATCAGGTGACGCGGAGGTTGA 1751
Db 1681 CCAGTATGTCGAGCTCGTCCATCGCTCAGTGGCCATCAGGTGACGCGGAGGTTGA 1740
QY 1752 GCGCCCTGAGGACCAAGATGACAAGATGTTGGGGAATTTCAGTGGCACTGGAGCGCTTC 1811
Db 1741 GCGCCCTGAGGACCAAGATGACAAGATGTTGGGGAATTTCAGTGGCACTGGAGCGCTTC 1800
QY 1812 CGACTACTCAGCGGCGCAACCCATTAAAGTGACATCGGTGCTACATCTTAGGAGACGA 1871
Db 1801 CGACTACTCAGCGGCGCAACCCATTAAAGTGACATCGGTGCTACATCTTAGGAGACGA 1860
QY 1872 CAGCTCAGTGTGACTGGACCTGTATCAAGTTCCTCGAGGCTGGAAAGACCAAGCT 1931
Db 1861 CAGCTCAGTGTGACTGGACCTGTATCAAGTTCCTCGAGGCTGGAAAGACCAAGCT 1920
QY 1932 GCACATCCAGCACAGATTGAACCTCGAGAACCAANTTAAGAACCTGAGGAGTCCG 1991
Db 1921 GCACATCCAGCACAGATTGAACCTCGAGAACCAANTTAAGAACCTGAGGAGTCCG 1980
QY 1992 AGGTCCTGAAGAAAAAGCGGCGCAGAGAAATGTGACTGTCAAAAATCAGCTACCCAC 2051

Db 1981 AGGTCCTGAAGAAAAAGCGGCGCAGAGAAATGTGACTGTCAAAAATCAGCTACCCAC 2040
QY 2052 CCAGCACAAAGGCGGCTCAAGCACAGAGGCTCCAGTCTGCATCCCTTCAGGAAGGCGCT 2111
Db 2041 CCAGCACAAAGGCGGCTCAAGCACAGAGGCTCCAGTCTGCATCCCTTCAGGAAGGCGCT 2100
QY 2112 GCAAGAGAAAGCAAGAGTGTGGCTGTGGGGAGCAGAAAGCGCAAGAACTCCGCGAA 2171
Db 2101 GCAAGAGAAAGCAAGAGTGTGGCTGTGGGGAGCAGAAAGCGCAAGAACTCCGCGAA 2160
QY 2172 GCTGCTCAAGCGGCTCGAGAACCAACAGCAGCTGCAGATGCCAGGCTCACTGCTTCAC 2231
Db 2161 GCTGCTCAAGCGGCTCGAGAACCAACAGCAGCTGCAGATGCCAGGCTCACTGCTTCAC 2220
QY 2232 CCAGCACAAACAGCACTGGCAGACGGCGGCTTCTCGACACTGGGCGCTTCTGTGCTG 2291
Db 2221 CCAGCACAAACAGCACTGGCAGACGGCGGCTTCTCGACACTGGGCGCTTCTGTGCTG 2280
QY 2292 CACAGGCGCAACATTAACAGCTACTGTGTGATGAGGACCATCAATGAGACTCAAAAT 2351
Db 2281 CACAGGCGCAACATTAACAGCTACTGTGTGATGAGGACCATCAATGAGACTCAAAAT 2340
QY 2352 CTTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTCTCAACACAGACCCCTA 2411
Db 2341 CTTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTCTCAACACAGACCCCTA 2400
QY 2412 CCAGCTGATGAATGCAAGTGAACACACTGCAAGGGATGCTCAACAGCTACAGTACA 2471
Db 2401 CCAGCTGATGAATGCAAGTGAACACACTGCAAGGGATGCTCAACAGCTACAGTACA 2460
QY 2472 GCTCATGGAGCTGAGGAGCTGCAAGGGTTTACAGCAGTGTAAACCCCGGACTCGAAACAT 2531
Db 2461 GCTCATGGAGCTGAGGAGCTGCAAGGGTTTACAGCAGTGTAAACCCCGGACTCGAAACAT 2520
QY 2532 GGAACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGTTCAGCGCTGAAA 2591
Db 2521 GGAACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGTTCAGCGCTGAAA 2580
QY 2592 GTGGCCAGAAATGAGAGAGCTTCTTCCAAATCAGTGGGACAACTGTGGGAAAGGCTGGA 2651
Db 2581 GTGGCCAGAAATGAGAGAGCTTCTTCCAAATCAGTGGGACAACTGTGGGAAAGGCTGGA 2640
QY 2652 AGGTTAAGAAAACACAGAGGCTGGAACCTTCCAAAACATAGAGGATCACTGACTGCACAG 2711
Db 2641 AGGTTAAGAAAACACAGAGGCTGGAACCTTCCAAAACATAGAGGATCACTGACTGCACAG 2700
QY 2712 GCATGAAAAACATGTGGGTGATTTCCAGCAGACCTGTGTCTATTTGGCCAGAGGCTGA 2771
Db 2701 GCATGAAAAACATGTGGGTGATTTCCAGCAGACCTGTGTCTATTTGGCCAGAGGCTGA 2760
QY 2772 GAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATTAACAGCAGAGCA 2831
Db 2761 GAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATTAACAGCAGAGCA 2820
QY 2832 GAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTCTTTGGATTTATACCTCACCAG 2891
Db 2821 GAGATAACTTTCAGGAAGTCCATTTTGGCCCTGCTTTTCTTTGGATTTATACCTCACCAG 2880
QY 2892 CTGCAAAAATGCAATTTTCTGATCAAAAAGTCAACCACTAACCCCTCCCGCAGAGCTCA 2951
Db 2881 CTGCAAAAATGCAATTTTCTGATCAAAAAGTCAACCACTAACCCCTCCCGCAGAGCTCA 2940
QY 2952 CAAAGGAAAACCGAGAGAGCGAGCGAGAGATTTCTTTGGAAAATTTCTCCCAAGGGCGA 3011
Db 2941 CAAAGGAAAACCGAGAGAGCGAGCGAGAGATTTCTTTGGAAAATTTCTCCCAAGGGCGA 3000
QY 3012 AAGTCATTTGGAAATTTTAAATCAATAGGGGAAAAGCAGTCTGTTCTAAATCTCTTATTC 3071
Db 3001 AAGTCATTTGGAAATTTTAAATCAATAGGGGAAAAGCAGTCTGTTCTAAATCTCTTATTC 3060
QY 3072 TTTTGGTTTGTCAAAAAGGAACTTAAGAGCAGGACAGAGGCAACCTGAGAGGCTGA 3131
Db 3061 TTTTGGTTTGTCAAAAAGGAACTTAAGAGCAGGACAGAGGCAACCTGAGAGGCTGA 3120

QY 1021 GGCTGTGTAAGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCGCTTCTACGTC 1080
Db 1583 GGCTGTGTAAGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCGCTTCTACGTC 1642
QY 1081 AGGGGCCCCAACGCTGGAGCGCGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG 1140
Db 1643 AGGGGCCCCAACGCTGGAGCGCGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG 1702
QY 1141 GCCCCCAACCTCTCTGGACATTGAGGCTGGACATACCTCGGATATGACGGGAATCC 1200
Db 1703 GCCCCCAACCTCTCTGGACATTGAGGCTGGACATACCTCGGATATGAGGGAATCC 1762
QY 1201 ATCTCAAGCTGTCTGGACAGCGCGCGGTGATTCGGTTTTCACCTGAAAGAGATG 1260
Db 1763 ATCTCAAGCTGTCTGGACAGCGCGCGGTGATTCGGTTTTCACCTGAAAGAGATG 1822
QY 1261 AGGTCTGGCGGACTCCTTCTTGGTGAGAGGCAAGCTGTACACAGAGAGACAAT 1320
Db 1823 AGGTCTGGCGGACTCCTTCTTGGTGAGAGGCAAGCTGTACACAGAGAGACAAT 1882
QY 1321 GACAGGTGAGCCCGAGGAGGAACTTCTGCCMAGTACAGCGTGTGAAGACCTG 1380
Db 1883 GACAGGTGAGCCCGAGGAGGAACTTCTGCCMAGTACAGCGTGTGAAGACCTG 1942
QY 1381 TGTCAAGCTGTAGTACAGACGCGCTGTGAGCAGCTGGGACAGAACTGGCAGTGTG 1440
Db 1943 TGTCAAGCTGTAGTACAGACGCGCTGTGAGCAGCTGGGACAGAACTGGCAGTGTG 2002
QY 1441 GAGACGCCACGGGGAAGCTGAAGCTGCATTAAGTGCAAGGCCCCATCGGCTGGCGGC 1500
Db 2003 GAGACGCCACGGGGAAGCTGAAGCTGCATTAAGTGCAAGGCCCCATCGGCTGGCGGC 2062
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGCGGACGGGACGAGCGCTGCAC 1560
Db 2063 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGCGGACGGGACGAGCGCTGCAC 2122
QY 1561 TGTGACAGCGGGACTACAAGCTCAGCTGGCGGACGCGGAAAAAATCTTTCAGAG 1620
Db 2123 TGTGACAGCGGGACTACAAGCTCAGCTGGCGGACGCGGAAAAAATCTTTCAGAG 2182
QY 1621 AAGTACAGGCCAGCTATGTCAGCTGCTCCATCGCTCAGTGGCCATCAGGTGGAC 1680
Db 2183 AAGTACAGGCCAGCTATGTCAGCTGCTCCATCGCTCAGTGGCCATCAGGTGGAC 2242
QY 1681 GGCAGGCTGTACCAAGCTAGGCTGGGTGATGCCGCCAGCCCGAACTCCACCAAGCG 1740
Db 2243 GGCAGGCTGTACCAAGCTAGGCTGGGTGATGCCGCCAGCCCGAACTCCACCAAGCG 2302
QY 1741 CACTGGCCAGGGCCCTGAGGACCAAGATGACAGGATGGTGGGAATTCAGTGGCACT 1800
Db 2303 CACTGGCCAGGGCCCTGAGGACCAAGATGACAGGATGGTGGGAATTCAGTGGCACT 2362
QY 1801 GGAGGCTCTCCGACTACTCAGCGGCCAACCCATTAAGTGGACATCGTGTACATC 1860
Db 2363 GGAGGCTCTCCGACTACTCAGCGGCCAACCCATTAAGTGGACATCGTGTACATC 2422
QY 1961 CTAGAGAACGACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGAGGCTGGAA 1920
Db 2423 CTAGAGAACGACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGAGGCTGGAA 2482
QY 1921 GACCAAGCTGCACATCGACACGAGATTGAACCTCGAGAACAAATTAAGAACCTG 1980
Db 2483 GACCAAGCTGCACATCGACACGAGATTGAACCTCGAGAACAAATTAAGAACCTG 2542
QY 1981 AGGGAAGTCCGAGGTCACTTGAAGAAAAAGCGGCCAGAGAAATGTCACTGTCACAAAATC 2040
Db 2543 AGGGAAGTCCGAGGTCACTTGAAGAAAAAGCGGCCAGAGAAATGTCACTGTCACAAAATC 2602
QY 2041 AGCTACCAACCCAGCAAAAGCGCCCTCAAGCAAGAGGCTCCAGTCTGCATCCTTTC 2100
Db 2603 AGCTACCAACCCAGCAAAAGCGCCCTCAAGCAAGAGGCTCCAGTCTGCATCCTTTC 2662

QY 2101 AGGAAGGCGCTGCAAGAGAAAGGACAAGGTGTGGCTGTGGCGGAGCAGAAAGCGCAAG 2160
Db 2663 AGGAAGGCGCTGCAAGAGAAAGGACAAGGTGTGGCTGTGGCGGAGCAGAAAGCGCAAG 2722
QY 2161 AAATCTCGCAAGCTGCTCAAGCGCTGCAGAACAAACACACGTCAGTCAGTCAGCGGCTC 2220
Db 2723 AAATCTCGCAAGCTGCTCAAGCGCTGCAGAACAAACACACGTCAGTCAGTCAGCGGCTC 2782
QY 2221 ACCTGCTTACCCACGACAAACACGACTGGCAGCGCGCTTCTTCTGGACACATCGGCGCT 2280
Db 2783 ACCTGCTTACCCACGACAAACACGACTGGCAGCGCGCTTCTTCTGGACACATCGGCGCT 2842
QY 2281 TTCTGTGCTGTCACAGCGCCCAACATAACACGTACTGTGTCATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGTCACAGCGCCCAACATAACACGTACTGTGTCATGAGGACCATCAATGAG 2902
QY 2341 ACTCACAATTTCTTCTTCTGTGAATTTGCAATTTGCAATTTCTAGAGTACTTTGATCTCAAC 2400
Db 2903 ACTCACAATTTCTTCTTCTGTGAATTTGCAATTTCTAGAGTACTTTGATCTCAAC 2962
QY 2401 ACAGACCCCTACAGCTGATGAATGACGTGAACACACTGACAGCGGATGCTCTCAACGAG 2460
Db 2963 ACAGACCCCTACAGCTGATGAATGACGTGAACACACTGACAGCGGATGCTCTCAACGAG 3022
QY 2461 CTACAGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGG 2520
Db 3023 CTACAGCTACAGCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGG 3082
QY 2521 ACTCGAAACATGGACCTTGGGACTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGT 2580
Db 3083 ACTCGAAACATGGACCTTGGGACTTAAAGATGGAGGAGCTATGAGCAATAACAGGAGT 3133
QY 2581 CAGGCTCGAAGTGGCCAGAAATGAGAGAGCTTCTTCAAAATCACTGGGACAACTGTGG 2640
Db 3134 CAGGCTCGAAGTGGCCAGAAATGAGAGAGCTTCTTCAAAATCACTGGGACAACTGTGG 3193
QY 2641 GAAGGCTGGGAAGGTTAAGAAACAAACAGAGGTGGACCTCCAAATAACATAGAGCATCACC 2700
Db 3194 GAAGGCTGGGAAGGTTAAGAAACAAACAGAGGTGGACCTCCAAATAACATAGAGCATCACC 3253
QY 2701 TGACTGACAGGCAATGAAAAAACAATGTGGTGAATTTCCAGCAGACCTGTGTATTTGGCC 2760
Db 3254 TGACTGACAGGCAATGAAAAAACAATGTGGTGAATTTCCAGCAGACCTGTGTATTTGGCC 3313
QY 2761 AGGAGGCTCAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3314 AGGAGGCTCAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
QY 2821 CAGCAGGACAGAGATAAATTTAGGAAGTCCATTTTGGCCCTGCTTTTGGATTG 2880
Db 3374 CAGCAGGACAGAGATAAATTTAGGAAGTCCATTTTGGCCCTGCTTTTGGATTG 3433
QY 2881 TACCTCAGCAGCTGCACAAATGCAATTTTTCGATATCAAAAGTCCACCTACCCCTCCC 2940
Db 3434 TACCTCAGCAGCTGCACAAATGCAATTTTTCGATATCAAAAGTCCACCTACCCCTCCC 3493
QY 2941 CCAGAGCTCACAAGGAAACCGAGAGCGGAGCGAGAGAGATTTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCACAAGGAAACCGAGAGCGGAGCGAGAGAGATTTTCTTGGAAATTTCT 3553
QY 3001 CCCAAGGCGGAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTCTTTCTAAA 3060
Db 3554 CCCAAGGCGGAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTCTTTCTAAA 3613
QY 3061 TCCTCTTATTTCTTTTGTGTCACAAAGAGGAACCTTAAGAGCAGGACAGAGGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTTGTGTCACAAAGAGGAACCTTAAGAGCAGGACAGAGGCAACGT 3673
QY 3121 GGAGAGGCTGAAACAGTGCAGAGAGCTTTGCAATGAGTCACTAGCACAACAAAGAGATGA 3180
Db 3674 GGAGAGGCTGAAACAGTGCAGAGAGCTTTGCAATGAGTCACTAGCACAACAAAGAGATGA 3733
QY 3181 CATTTACTAGCACTATAAACCCCTGCTTGGTGGCTCTGAGAAACTGCTTCACTGTATAT 3240

Db 3734 CATTACTAGCATATAAACCCCTGGTTCCTCTGAGAAACCTGCCCTTCATTGTATATAT 3793
Qy 3241 GTGACTATTACATGTATCAACATGGGAACTTTTAGGGGAACCTATAAGAAATCCCAA 3300
Db 3794 GTGACTATTACATGTATCAACATGGGAACTTTTAGGGGAACCTATAAGAAATCCCAA 3853
Qy 3301 TTTTCAGAGTGGTGGTGTCAATAAACGCTGTGGCCACTGTAAAGAAAAA 3353
Db 3854 TTTTCAGAGTGGTGGTGTCAATAAACGCTGTGGCCACTGTAAAGAAAAA 3906

RESULT 9

US-09-989-723-83
Sequence 83 Application US/09898723
Patent No. US20020072092A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C62
CURRENT APPLICATION NUMBER: US/09/989,723
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19

1643 AGGGCCCCAAGCTGGAGCGGCTGTCTGAATCCCACTGTCCTCAATTTGACCTG 1702
1141 GCCCCACATCTCTGGACATTTGCGAGCTGGACATACCTTGGGATATGGAACGGAAATCC 1200
1703 GCCCCACATCTCTGGACATTTGCGAGCTGGACATACCTTGGGATATGGAACGGAAATCC 1762
1201 ATCTTCAAGCTCTGGACACCGAGCGCGGTGAATCGGTTTCACTTCAAAAGAGATG 1260
1763 ATCTTCAAGCTCTGGACACCGAGCGCGGTGAATCGGTTTCACTTCAAAAGAGATG 1822
1261 AGGGTCTGGCGGACTCTCTTCTTGGAGAGAGCAAGCTGTCTACACAGAGAGACAAT 1320
1823 AGGGTCTGGCGGACTCTCTTCTTGGAGAGAGCAAGCTGTCTACACAGAGAGACAAT 1882
1321 GACAGAGTGGACGCCAGAGAGAGAACTTTCTGCCCAAGTACCAAGCTGTGAAGGACTG 1380
1893 GACAGAGTGGACGCCAGAGAGAGAACTTTCTGCCCAAGTACCAAGCTGTGAAGGACTG 1942
1381 TGTACAGCTGTCTGAGTACCAAGACGGCGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440
1943 TGTACAGCTGTCTGAGTACCAAGACGGCGTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2002
1441 GAGGACGCCACCGGGAAGCTGAAGCTGATTAAGTGCAAGGCGCCCATGCGGCTGGCGGC 1500
2003 GAGGACGCCACCGGGAAGCTGAAGCTGATTAAGTGCAAGGCGCCCATGCGGCTGGCGGC 2062
1501 AGCAGAGCCCTCTCAACTCTGTGCCCAAGTACTACGGGAGGCGGACGAGGCGCTGCACC 1560
2063 AGCAGAGCCCTCTCAACTCTGTGCCCAAGTACTACGGGAGGCGGACGAGGCGCTGCACC 2122
1561 TGTGACAGCGGGACTACAAGCTCAGCTGCGCGGACGCGCGGAAAAAATCTTCAAGAAG 1620
2123 TGTGACAGCGGGACTACAAGCTCAGCTGCGCGGACGCGCGGAAAAAATCTTCAAGAAG 2182
1621 AGTACAGGCGAGCTATGTCGAGCTGCTCCATCTCGCTCAGTGGGCGCATCAGGTGCAC 1680
2183 AGTACAGGCGAGCTATGTCGAGCTGCTCCATCTCGCTCAGTGGGCGCATCAGGTGCAC 2242
1681 GCGAGGCTTACCACTAGGCTGTGGTATGATCGCGCCAGCGCCGAAACCTCACCACGCGG 1740
2243 GCGAGGCTTACCACTAGGCTGTGGTATGATCGCGCCAGCGCCGAAACCTCACCACGCGG 2302
1741 CACTGGCCAGGGGCCCCCTGAGGACCAAGATGACAGGATGTTGGGACTTCAGTGGCACT 1800
2303 CACTGGCCAGGGGCCCCCTGAGGACCAAGATGACAGGATGTTGGGACTTCAGTGGCACT 2362
1801 GGAGGCTTCCCGACTACTCAGCGGCCAACCCCATTTAAAGTGACACATCGGTGCTATC 1860
2363 GGAGGCTTCCCGACTACTCAGCGGCCAACCCCATTTAAAGTGACACATCGGTGCTATC 2422
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGCTGGAAA 1920
2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAGGCTGGAAA 2482
1921 GACCAAGCTGCACATCGACACAGATTTGAACCTCTGAGAACAAATTAAGAACCTG 1980
2483 GACCAAGCTGCACATCGACACAGATTTGAACCTCTGAGAACAAATTAAGAACCTG 2542
1981 AGGGAAGTCCGAGGTCACTGAAAGAAAAGCGCGCAGAGAAATGTGACTGTCAAAAATC 2040
2543 AGGGAAGTCCGAGGTCACTGAAAGAAAAGCGCGCAGAGAAATGTGACTGTCAAAAATC 2602
2041 AGTTACACACCCAGCAAAAGGCGCTCAAGCACAGAGGCTCCAGTGTGATCTTTC 2100
2603 AGTTACACACCCAGCAAAAGGCGCTCAAGCACAGAGGCTCCAGTGTGATCTTTC 2662
2101 AGGAAGGCTTGCAGAGAGAGACAGAGTGTGGTGTGGGAGCAGAGCGCAGAGAG 2160
2663 AGGAAGGCTTGCAGAGAGAGACAGAGTGTGGTGTGGGAGCAGAGCGCAGAGAG 2722
2161 AAACCTCCGAAAGCTGTCTCAAGCGCTTCAGAAACACGACAGTGCAGCATGCCAGGCTC 2220

2723 AAACCTCCGAAAGCTGTCTCAAGCGCTTCAGAGAACACAGCACAGTGCAGCATGCCAGGCTC 2782
2221 ACCTGCTTCACCACGACAAACAGACACTGGCAGACGGCGCTTTCTTGACACTTGGGCTC 2280
2783 ACCTGCTTCACCACGACAAACAGACACTGGCAGACGGCGCTTTCTTGACACTTGGGCTC 2842
2281 TTTCTGCTCTGACACGAGCGCAACATATACAGTACTGTGTATGAGGACCATCAATGAG 2340
2843 TTTCTGCTCTGACACGAGCGCAACATATACAGTACTGTGTATGAGGACCATCAATGAG 2902
2341 ACTCACAATTTCTCTTCTGTGTGAATTTGCAATTTGGCTTCTTAGAGTACTTTGATCTCAAC 2400
2903 ACTCACAATTTCTCTTCTGTGTGAATTTGCAATTTGGCTTCTTAGAGTACTTTGATCTCAAC 2962
2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGAAGGATGCTCTCAACCAAG 2460
2963 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGGAAGGATGCTCTCAACCAAG 3022
2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTCAAGGGTTCAAGCAGTGTAAACCCCGG 2520
3023 CTACAGCTACAGCTCATGAGCTGAGGAGCTCAAGGGTTCAAGCAGTGTAAACCCCGG 3082
2521 ACTCGAAAATGGAACCTGGAATTAAGATGGAAGGATATGAGCAATACAGGCACTTT 2580
3083 ACTCGAAAATGGAACCTGGAATTAAGATGGAAGGATATGAGCAATACAGGCACTTT 3133
2581 CAGCGTCGAAAGTGGCCAGAAATGAGAGACCTTTCTTCCNAATCACTGGGACAACTGTGG 2640
3134 CAGCGTCGAAAGTGGCCAGAAATGAGAGACCTTTCTTCCNAATCACTGGGACAACTGTGG 3193
2641 GAAGGCTGGAAAGTTAAGAAAACAAACAGAGGTGGACCTCCAAAAACATAGAGGCACTACC 2700
3194 GAAGGCTGGAAAGTTAAGAAAACAAACAGAGGTGGACCTCCAAAAACATAGAGGCACTACC 3253
2701 TGACTGCACAGCAATGAAAAACCAATGTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCC 2760
3254 TGACTGCACAGCAATGAAAAACCAATGTGGGTGATTTCCAGCAGACCTGTGTCTATTGGCC 3313
2761 AGAGGCTTGAAGAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAC 2820
3314 AGAGGCTTGAAGAAAGCAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAC 3373
2821 CAGCAGGAGCAGAGATAAATCTCAGAAAGTCCATTTTTCCTGCTTTTGGTATTA 2880
3374 CAGCAGGAGCAGAGATAAATCTCAGAAAGTCCATTTTTCCTGCTTTTGGTATTA 3433
2881 TACCTCAGCAGCTGACAAATGCAATTTTTCGATCAAAAAGTCAACCACTAACCTCTCC 2940
3434 TACCTCAGCAGCTGACAAATGCAATTTTTCGATCAAAAAGTCAACCACTAACCTCTCC 3493
2941 CCAGAGGCTCAAAAAGGAAAACCGAGAGAGCGGAGAGATTTCTTGGAAATTTCT 3000
3494 CCAGAGGCTCAAAAAGGAAAACCGAGAGAGCGGAGAGATTTCTTGGAAATTTCT 3553
3001 CCAGAGGCGAAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3060
3554 CCAGAGGCGAAAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3613
3061 TCCCTTATTTCTTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGAGCAGGCAACGT 3120
3614 TCCCTTATTTCTTTTGGTTTGTCAAAAAGAGGAACTAAGAGCAGGAGCAGGCAACGT 3673
3121 GGAGGCTGAAAAACAGTGCAGAGACGTTTGCAATATGATCAGTAGCAAAAAGAGATGA 3180
3674 GGAGGCTGAAAAACAGTGCAGAGACGTTTGCAATATGATCAGTAGCAAAAAGAGATGA 3733
3181 CATTTACCTAGCAGCTATAAACCCTGTTGCTCTGAGAAACCTGCTTCAATGTATATAT 3240
3734 CATTTACCTAGCAGCTATAAACCCTGTTGCTCTGAGAAACCTGCTTCAATGTATATAT 3793
3241 GTGAGCTATTTTACATGATCAACATGGGAACTTTTATGGGGAACTTAATAAGAAATCCCAA 3300
3794 GTGAGCTATTTTACATGATCAACATGGGAACTTTTATGGGGAACTTAATAAGAAATCCCAA 3853

QY 3301 TTTTCAGGAGTGGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3353
Db 3854 TTTTCAGGAGTGGTGTGTCAATAAAGCTCTGTGGCCAGTGTAAAGAAAA 3906

RESULT 10

US-09-989-279-83
/ Sequence 83, Application US/09989279
/ Patent No. US20020072496A1
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Kljavin, Ivar J.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: P2730P1C56
/ CURRENT APPLICATION NUMBER: US/09/989,279
/ CURRENT FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/049787
/ PRIOR FILING DATE: 1997-06-16
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/065186
/ PRIOR FILING DATE: 1997-11-12
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066770
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/075945
/ PRIOR FILING DATE: 1998-02-25
/ PRIOR APPLICATION NUMBER: 60/078910
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/084600
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/087106
/ PRIOR FILING DATE: 1998-05-28
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087609
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087759
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087827
/ PRIOR FILING DATE: 1998-06-03
/ PRIOR APPLICATION NUMBER: 60/088021
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088025
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088026

/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088028
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088029
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088030
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088033
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088326
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088167
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088202
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088212
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088217
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088655
/ PRIOR FILING DATE: 1998-06-09
/ PRIOR APPLICATION NUMBER: 60/088734
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088738
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088742
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088810
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088824
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088826
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088858
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088861
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088876
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/089105
/ PRIOR FILING DATE: 1998-06-12
/ PRIOR APPLICATION NUMBER: 60/089440
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089512
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089514
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089532
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089538
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089598
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089599
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089600
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089653
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089907
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089908
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089947
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089948
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089952
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/090246
/ PRIOR FILING DATE: 1998-06-22

RESULT 11
US-09-989-727-83
Sequence 83, Application US/09899727
Patent No. US20020072497A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Garber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C65
CURRENT APPLICATION NUMBER: US/09/989,727
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23

Db 1823 AGGCTCTGGCGGAGCTCTCTTCTTGGTGAGAGGCAAGCTGTCTACACAGAGAGACAAT 1882
Qy 1321 GACAAGGTGACAGCGCCAGGAGGAGAACTTTCTGCCCAAGTACCAGAGGTGTGAAGGACCTG 1380
Db 1883 GACAAGGTGACAGCGCCAGGAGGAGAACTTTCTGCCCAAGTACCAGAGGTGTGAAGGACCTG 1942
Qy 1381 TGTGAGCGTGTGAGTACCAAGAGCGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440
Db 1943 TGTGAGCGTGTGAGTACCAAGAGCGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 2002
Qy 1441 GAGGACGCCAGCGGGAAGCTGAAGCTCATAGTGAAGGCGCCCATGGCGCTGGCGCGC 1500
Db 2003 GAGGACGCCAGCGGGAAGCTGAAGCTCATAGTGAAGGCGCCCATGGCGCTGGCGCGC 2062
Qy 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGAGGCGGAGCGGCTGCACC 1560
Db 2063 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTACGGGAGGCGGAGCGGCTGCACC 2122
Qy 1561 TGTGACAGCGGGAAGTCAAGCTCAGCTGGCGGAGCGCGGAAACCTTTCRAAGAG 1620
Db 2123 TGTGACAGCGGGAAGTCAAGCTCAGCTGGCGGAGCGCGGAAACCTTTCRAAGAG 2182
Qy 1621 AAGTACAGGCGAGCTATGTCGCGAGTGTCTCCATCGCTCAGTGGCCATCGAGGTGGAC 1680
Db 2183 AAGTACAGGCGAGCTATGTCGCGAGTGTCTCCATCGCTCAGTGGCCATCGAGGTGGAC 2242
Qy 1681 GGCAGGTGTACCAAGTGTGCTGGTGTGATGCGCGCCAGCGCCGAAACCTTCACCAAGCG 1740
Db 2243 GGCAGGTGTACCAAGTGTGCTGGTGTGATGCGCGCCAGCGCCGAAACCTTCACCAAGCG 2302
Qy 1741 CACTGGCCAGGGCCCTGAGGACCAAGTGAAGGATGGTGGGACTTCAGTGGCACT 1800
Db 2303 CACTGGCCAGGGCCCTGAGGACCAAGTGAAGGATGGTGGGACTTCAGTGGCACT 2362
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAAGTGAACATCGGTGTACATC 1860
Db 2363 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAAGTGAACATCGGTGTACATC 2422
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACTGGACCTGTACAGTCCCTGCGGCGCTGGAAA 1920
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACTGGACCTGTACAGTCCCTGCGGCGCTGGAAA 2482
Qy 1921 GACCACAAGTGTGCATCGACACAGATTTGAAACCTCTGACAGAACAAATTAAGAACCTG 1980
Db 2483 GACCACAAGTGTGCATCGACACAGATTTGAAACCTCTGACAGAACAAATTAAGAACCTG 2542
Qy 1981 AGGGAAGTCCGAGTCACTGAAGAAAGCGCCAGAGAAATGTGACTGTCACAAAATC 2040
Db 2543 AGGGAAGTCCGAGTCACTGAAGAAAGCGCCAGAGAAATGTGACTGTCACAAAATC 2602
Qy 2041 AGCTACACACCCAGACAAAGCGCCCTCAAGCACAGAGGCTCCAGTCTGCATCCCTTTC 2100
Db 2603 AGCTACACACCCAGACAAAGCGCCCTCAAGCACAGAGGCTCCAGTCTGCATCCCTTTC 2662
Qy 2101 AGGAAGGCGCTGCAAGAGAGGCAAGGTGTGCTGTGGGAGCAGAGCGCAAGAAG 2160
Db 2663 AGGAAGGCGCTGCAAGAGAGGCAAGGTGTGCTGTGGGAGCAGAGCGCAAGAAG 2722
Qy 2161 AAATCCCGAGCTGTCTCAAGCGCTGCAGAACACAGACAGTGCAGATGCCAGGCTC 2220
Db 2723 AAATCCCGAGCTGTCTCAAGCGCTGCAGAACACAGACAGTGCAGATGCCAGGCTC 2782
Qy 2221 ACGTGCTTCAACCCAGCAACACGACTGGCAGACGGCGCTTTCTGGACACTGGGGCT 2280
Db 2783 ACGTGCTTCAACCCAGCAACACGACTGGCAGACGGCGCTTTCTGGACACTGGGGCT 2842
Qy 2281 TTCTGTGCTGTGACAGCGGCCCAACATPAAACAGTATCTGGTGCATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGTGACAGCGGCCCAACATPAAACAGTATCTGGTGCATGAGGACCATCAATGAG 2902
Qy 2341 ACTCACAATTTCTCTCTGTGATTTGCAACTGGCTTCTCAGAGTACTTTGATCTCAAC 2400

Db 2903 ACTCACAATTTCTCTCTGTGAAATTTGCACACTGGCTTCTTAGAGTACTTTGATCTCAAC 2962
Qy 2401 ACAGACCCCTACCAAGCTGATGAATGCACTGAGACAGGATGTCTCAACCCAG 2460
Db 2963 ACAGACCCCTACCAAGCTGATGAATGCACTGAGACAGGATGTCTCAACCCAG 3022
Qy 2461 CTACACGTACAGCTCATGAGCTGAGAGCTGCAAGGGTTACAAGCGAGTGTAAACCCCGG 2520
Db 3023 CTACACGTACAGCTCATGAGCTGAGAGCTGCAAGGGTTACAAGCGAGTGTAAACCCCGG 3082
Qy 2521 ACTCGAATCATGAGACCTTGAAGATGAGAGAGCTATGAGCAATACAGAGCAGTTT 2580
Db 3083 ACTCGAATCATGAGACCTTGAAGATGAGAGAGCTATGAGCAATACAGAGCAGTTT 3133
Qy 2581 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTTCTCCAAATCACTGGGCAACTGTGG 2640
Db 3134 CAGCGTCCAAAGTGGCCAGAAATGAAGAGACCTTTCTCCAAATCACTGGGCAACTGTGG 3193
Qy 2641 GAAGCTGGGAGGTGAAGAAACACAGAGGTGACCTTCCAAACATAGAGGCACTCACC 2700
Db 3194 GAAGCTGGGAGGTGAAGAAACACAGAGGTGACCTTCCAAACATAGAGGCACTCACC 3253
Qy 2701 TGACTGCAACGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 2760
Db 3254 TGACTGCAACGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 3313
Qy 2761 AGGAGGCTGAGAAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3314 AGGAGGCTGAGAAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
Qy 2821 CAGCAGGAGCAGAGATTAATTTGAGAAAGTCAATTTTGGCCCTCTTTTGGGATTA 2880
Db 3374 CAGCAGGAGCAGAGATTAATTTGAGAAAGTCAATTTTGGCCCTCTTTTGGGATTA 3433
Qy 2881 TACTCAGCAGCTGCACAAAATGCAATTTTTCGATCAAAAAGTCAACACTAACCTTCCC 2940
Db 3434 TACTCAGCAGCTGCACAAAATGCAATTTTTCGATCAAAAAGTCAACACTAACCTTCCC 3493
Qy 2941 CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553
Qy 3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTCTTCTAAA 3060
Db 3554 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTCTTCTAAA 3613
Qy 3061 TCCTCTTATTTCTTTGGTTGTCTCAAAAGAGGAACTAAGAGCAGGACAGAGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTGGTTGTCTCAAAAGAGGAACTAAGAGCAGGACAGAGCAACGT 3673
Qy 3121 GGAGAGGCTGAAAAACAGTGCAGAGACGTTTGAATGAGTCAGTCAGTGCACAAAGAGATGA 3180
Db 3674 GGAGAGGCTGAAAAACAGTGCAGAGACGTTTGAATGAGTCAGTCAGTGCACAAAGAGATGA 3733
Qy 3181 CATTTACCTAGCATATAAACCTTGGTTGCTCTGAAAGAACTGCGCTTCTCATTTGATATAT 3240
Db 3734 CATTTACCTAGCATATAAACCTTGGTTGCTCTGAAAGAACTGCGCTTCTCATTTGATATAT 3793
Qy 3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACTTAATAGAAATCCCAA 3300
Db 3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACTTAATAGAAATCCCAA 3853
Qy 3301 TTTTCAGGAGTGGTGTCTCAATAAAGCTCTGTGGCCAGTGTAAAAAGAAAAA 3353
Db 3854 TTTTCAGGAGTGGTGTCTCAATAAAGCTCTGTGGCCAGTGTAAAAAGAAAAA 3906

RESULT 12

US-09-989-731-83
; Sequence 83, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24

;	PRIOR APPLICATION NUMBER: 60/090435																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
---	-------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

1381 TGTCAGCGTGTGAGTACAGACGGCGTGTGAGCAGTGGGACAGAGTGGCGAGTGTGTG 1440
1943 TGTCAGCGTGTGAGTACAGACGGCGTGTGAGCAGTGGGACAGAGTGGCGAGTGTGTG 2002
1441 GAGGACGCCACGGGGAAGCTGAAGCTGCAATAAGTGAAGGCCGCCATCGCGCTGGCGGC 1500
2003 GAGGACGCCACGGGGAAGCTGAAGCTGCAATAAGTGAAGGCCGCCATCGCGCTGGCGGC 2062
1501 AGCAGAGCCCTCTCCAACTCGTGGCCCAAGTACTAGGGCAGGACGAGGCGCTGCACC 1560
2063 AGCAGAGCCCTCTCCAACTCGTGGCCCAAGTACTAGGGCAGGACGAGGCGCTGCACC 2122
1561 TGTCAGCGGGGACTACAAGCTCAGCCTGGCGGACGCGCGGAAAGAACTCTTCAAGAG 1620
2123 TGTCAGCGGGGACTACAAGCTCAGCCTGGCGGACGCGCGGAAAGAACTCTTCAAGAG 2182
1621 AAGTACAGGCGAGTATGTCGCGAGTCCCTCCATCCGCTAGTGGCCATCGAGTGGAC 1680
2183 AAGTACAGGCGAGTATGTCGCGAGTCCCTCCATCCGCTAGTGGCCATCGAGTGGAC 2242
1681 GGCAGGGTGTACACAGTAGGCTGGGTGATGCGGCCCGAGCCCGGAACTCACCAGCGG 1740
2243 GGCAGGGTGTACACAGTAGGCTGGGTGATGCGGCCCGAGCCCGGAACTCACCAGCGG 2302
1741 CACTGCGCAGGGCCCTGAGGACCGAGATGACAGGATGAGTGGGACCTTCAAGTGGCACT 1800
2303 CACTGCGCAGGGCCCTGAGGACCGAGATGACAGGATGAGTGGGACCTTCAAGTGGCACT 2362
1801 GAGGGCTTCCCGACTACTCAGCGGCCACCCATTAAGTGCACATCGGTGTACATC 1860
2363 GAGGGCTTCCCGACTACTCAGCGGCCACCCATTAAGTGCACATCGGTGTACATC 2422
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTCGAGCCCTGGAAA 1920
2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTCGAGCCCTGGAAA 2482
1921 GACCAAGCTGCACATCGACACGAGATTGAACCTCGCAGAACAAATTAAGAACCTG 1980
2483 GACCAAGCTGCACATCGACACGAGATTGAACCTCGCAGAACAAATTAAGAACCTG 2542
1981 AGGGAAGTCCGAGGTCACTGAAGAAAAGCGGCGCAGAGAAATGTGACTGTCACAAAATC 2040
2543 AGGGAAGTCCGAGGTCACTGAAGAAAAGCGGCGCAGAGAAATGTGACTGTCACAAAATC 2602
2041 AGCTACACACCCAGCACAAGCGCGCTCAGCAGCAGAGGCTCAGTCTGATCTTTC 2100
2603 AGCTACACACCCAGCACAAGCGCGCTCAGCAGCAGAGGCTCAGTCTGATCTTTC 2662
2101 AGGAAGGCGCTGCAAGAGAGGACAGAGTGTGGCTGTTCGGGAGCAGAGCGCAAGAG 2160
2663 AGGAAGGCGCTGCAAGAGAGGACAGAGTGTGGCTGTTCGGGAGCAGAGCGCAAGAG 2722
2161 AAACCTCGGAAGTGTCTAAGCGCTTCAGAACAAAGCAGCAGTGTGAGCAGATGCGCGCTC 2220
2723 AAACCTCGGAAGTGTCTAAGCGCTTCAGAACAAAGCAGCAGTGTGAGCAGATGCGCGCTC 2782
2221 AGTGTCTTACCCACGACACAGCAGTGTGAGCAGCGCGCTTCTGAGCAGTGTGGCGCT 2280
2783 AGTGTCTTACCCACGACACAGCAGTGTGAGCAGCGCGCTTCTGAGCAGTGTGGCGCT 2842
2281 TTCTGTGCTGCACAGCGCCCAACATAACAGTACTGTGTGATGAGGACCATCAATGAG 2340
2843 TTCTGTGCTGCACAGCGCCCAACATAACAGTACTGTGTGATGAGGACCATCAATGAG 2902
2341 ACTCAAAATTTCTCTTCTGTAATTTGCAACTGCTTCTAGTACTTTGATCTCAAC 2400
2903 ACTCAAAATTTCTCTTCTGTAATTTGCAACTGCTTCTAGTACTTTGATCTCAAC 2962
2401 ACAGACCCCTTACAGCTGATGATGAGTGAAGACACACTGGACAGGGGATGTCTCAACAG 2460
2963 ACAGACCCCTTACAGCTGATGATGAGTGAAGACACACTGGACAGGGGATGTCTCAACAG 3022

2461 CTACAGTACAGCTCATGAGCTGAGAGCTGCAAGCGGTTCAAGCAGTGTAAACCCCGG 2520
3023 CTACAGTACAGCTCATGAGCTGAGAGCTGCAAGCGGTTCAAGCAGTGTAAACCCCGG 3082
2521 ACTCGAAAACATGACCTGGGACTTAAAGATGAGGAGCTATGAGCAATACAGGCAAGTTT 2580
3083 ACTCGAAAACATGACCTGGGACTTAAAGATGAGGAGCTATGAGCAATACAGGCAAGTTT 3133
2581 CAGCGTCGAAGTGGCGAGAAATGAAGAGACCTTCTCCAAATCACTGGGACAACTGTGG 2640
3134 CAGCGTCGAAGTGGCGAGAAATGAAGAGACCTTCTCCAAATCACTGGGACAACTGTGG 3193
2641 GAAGGCTGGGAAGTTAAGAAAACACAGAGTGGACCTCCAAAACATAGAGCATCAC 2700
3194 GAAGGCTGGGAAGTTAAGAAAACACAGAGTGGACCTCCAAAACATAGAGCATCAC 3253
2701 TGACTGCACAGGCAATGAAGAAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 2760
3254 TGACTGCACAGGCAATGAAGAAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 3313
2761 AGGAGGCTGGGAAGTTAAGAAAACACAGAGTGGACCTCCAAAACATAGAGCATCAC 2820
3314 AGGAGGCTGGGAAGTTAAGAAAACACAGAGTGGACCTCCAAAACATAGAGCATCAC 3373
2821 CAGCAGGACGAGATTAACCTTCAAGAGTCCATTTTGGCCCTGTCTTTGCTTTGATTA 2880
3374 CAGCAGGACGAGATTAACCTTCAAGAGTCCATTTTGGCCCTGTCTTTGCTTTGATTA 3433
2881 TACTCACCAGCTGCACAAAATGCAATTTTGGTATCAAAAAGTCAACACTAACCTCC 2940
3434 TACTCACCAGCTGCACAAAATGCAATTTTGGTATCAAAAAGTCAACACTAACCTCC 3493
2941 CCAGAGCTCACAAGGAAAACCGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3000
3494 CCAGAGCTCACAAGGAAAACCGAGAGCGAGGAGAGATTTCTTGGAAATTTCT 3553
3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAA 3060
3554 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAGCAGTCTTCTTAA 3613
3061 TCCTCTTATTTCTTTGGTTGTCAAAAAGGAACTAAGAAGCAGGACAGAGCAACGT 3120
3614 TCCTCTTATTTCTTTGGTTGTCAAAAAGGAACTAAGAAGCAGGACAGAGCAACGT 3673
3121 GGAGGCTGAAAACAGTGCAGAGAGTGTGCAATGAGTCAAGTACAGAGAGATGA 3180
3674 GGAGGCTGAAAACAGTGCAGAGAGTGTGCAATGAGTCAAGTACAGAGAGATGA 3733
3181 CATTTACCTAGCACTATAAACCTTGTCTCTGAGAAACTGCTTCATTTGATATAT 3240
3734 CATTTACCTAGCACTATAAACCTTGTCTCTGAGAAACTGCTTCATTTGATATAT 3793
3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTTAATAAGAAATCCCAA 3300
3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTTAATAAGAAATCCCAA 3853
3301 TTTTACAGGAGTGTGTGTCAATTAACAGCTCTGTGGCCAGTGTAAAGAAAAA 3353
3854 TTTTACAGGAGTGTGTGTCAATTAACAGCTCTGTGGCCAGTGTAAAGAAAAA 3906

RESULT 13

US-09-989-732-83
; Sequence 83, Application US/09989732
; Patent No. US20020123463A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoli, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C57
CURRENT APPLICATION NUMBER: US/09/989,732
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-15
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24

1	Prior Application Number: 60/090472	QY	361	ACCTACACCAACAACTAGAACTGCTCCTCGCCCTCTGTCAGGCAACAGACGAGAGCGGC	420
2	Prior Filing Date: 1998-06-24	Db	923	ACCTACACCAACAACTAGAACTGCTCCTCGCCCTCTGTCAGGCAACAGACGAGAGCGGC	982
3	Prior Application Number: 60/090535	QY	421	ACCTTTGCGGTGTACTCTCAATAGCACTGGCTACCGGACAGCTTTCTTTCGGGAAGTATCTT	480
4	Prior Filing Date: 1998-06-24	Db	983	ACCTTTGCGGTGTACTCTCAATAGCACTGGCTACCGGACAGCTTTCTTTCGGGAAGTATCTT	1042
5	Prior Application Number: 60/090540	QY	481	ATGTAATACAAACGGCTCTTACGTGCCACCGGCTCGAAGAGTGGGTGGAATCTCTTAA	540
6	Prior Filing Date: 1998-06-24	Db	1043	ATGTAATACAAACGGCTCTTACGTGCCACCGGCTCGAAGAGTGGGTGGAATCTCTTAA	1102
7	Prior Application Number: 60/090542	QY	541	AACTCCCGCTTTTATAAATACTACAGCTGTGTGCGAAGCGGCTGAAAGAAAGCAAGCGCTCC	600
8	Prior Filing Date: 1998-06-25	Db	1103	AACTCCCGCTTTTATAAATACTACAGCTGTGTGCGAAGCGGCTGAAAGAAAGCAAGCGCTCC	1162
9	Prior Application Number: 60/090557	QY	601	GACTACTCCAAAGATTTACCTCACAGACCTCATCAACATGACAGCGTGTGCTTTCCGC	660
10	Prior Filing Date: 1998-06-25	Db	1163	GACTACTCCAAAGATTTACCTCACAGACCTCATCAACATGACAGCGTGTGCTTTCCGC	1222
11	Prior Application Number: 60/090576	QY	661	AGTCCAAAGATGTACCCGACAGGCGAGTCTCTCATGTGTATCAGCCATCAGCCCTCC	720
12	Prior Filing Date: 1998-06-25	Db	1223	AGTCCAAAGATGTACCCGACAGGCGAGTCTCTCATGTGTATCAGCCATCAGCCCTCC	1282
13	Prior Application Number: 60/090696	QY	721	CAGGCGCTGAGGATTCAGCCCAACATATTCAGGCTCTTCCCAACGCAATCTCAGAC	780
14	Prior Filing Date: 1998-06-26	Db	1283	CAGGCGCTGAGGATTCAGCCCAACATATTCAGGCTCTTCCCAACGCAATCTCAGAC	1342
15	Prior Application Number: 60/090863	QY	781	ATCACCGGAGCTTACAACTACGCGCCCAACCGCGAACAACACTGGATCATGCGCTACAG	840
16	Prior Filing Date: 1998-07-01	Db	1343	ATCACCGGAGCTTACAACTACGCGCCCAACCGCGAACAACACTGGATCATGCGCTACAG	1402
17	Prior Application Number: 60/091360	QY	841	GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAGGCTTGAG	900
18	Prior Filing Date: 1998-07-02	Db	1403	GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAGGCTTGAG	1462
19	Prior Application Number: 60/091478	QY	901	ACCTCATGTCTGGTGAAGCTTCCATGAGAGGATTTACCAACATGCTGGTTGAGACGGGC	960
20	Prior Filing Date: 1998-07-02	Db	1463	ACCTCATGTCTGGTGAAGCTTCCATGAGAGGATTTACCAACATGCTGGTTGAGACGGGC	1522
21	Prior Application Number: 60/091633	QY	961	GAGCTGGACAAACGCTACATCGTATACACCGCGACCAACGCTTACCATCGGCCAGTTT	1020
22	Prior Filing Date: 1998-07-02	Db	1523	GAGCTGGACAAACGCTACATCGTATACACCGCGACCAACGCTTACCATCGGCCAGTTT	1582
23	Prior Application Number: 60/091978	QY	1021	GGCTGGTGAAGGGAATCCCATGATATGATTTGACATCAGGCTCCGTTCTACGTCG	1080
24	Prior Filing Date: 1998-07-02	Db	1583	GGCTGGTGAAGGGAATCCCATGATATGATTTGACATCAGGCTCCGTTCTACGTCG	1642
25	Prior Application Number: 60/091982	QY	1081	AGGGGCCCCAACGTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG	1140
26	Prior Filing Date: 1998-07-07	Db	1643	AGGGGCCCCAACGTGGAAGCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG	1702
27	Prior Application Number: 60/092182	QY	1141	GGCCCAACCATCTGAGACATTCAGCGCTGACATACCTGCGGATATGACCGGGAATCC	1200
28	Prior Filing Date: 1998-07-09	Db	1703	GGCCCAACCATCTGAGACATTCAGCGCTGACATACCTGCGGATATGACCGGGAATCC	1762
29	Prior Application Number: 60/091519	QY	1201	ATCCTCAAGCTGTGACACGAGCGCGCGGTGAATCGGTTTCACTTGAAAAAGAGATG	1260
30	Prior Filing Date: 1998-07-02	Db	1763	ATCCTCAAGCTGTGACACGAGCGCGCGGTGAATCGGTTTCACTTGAAAAAGAGATG	1822
31	Prior Application Number: 60/091626	QY	1261	AGGGTCTGGGGGATCTCTTCTTGTGGAGAGAGGCAAGCTCTACACAGAGACAAAT	1320
32	Prior Filing Date: 1998-07-02	Db	1823	AGGGTCTGGGGGATCTCTTCTTGTGGAGAGAGGCAAGCTCTACACAGAGACAAAT	1882
33	Prior Application Number: 60/091633	QY	1321	GACCAAGTGGACCGCCAGGAGGAGAACTTTCTGCCCAAGTACAGCTGTGAGGACCTG	1380
34	Prior Filing Date: 1998-07-01	Db	1883	GACCAAGTGGACCGCCAGGAGGAGAACTTTCTGCCCAAGTACAGCTGTGAGGACCTG	1942
35	Prior Application Number: 60/091978	QY	1381	TGTACAGCTGTGAGTACAGACGCGCTGTGAGAGCTGGGACAGAGTGGCAGTGTGTG	1440
36	Prior Filing Date: 1998-07-07	Db	1943	TGTACAGCTGTGAGTACAGACGCGCTGTGAGAGCTGGGACAGAGTGGCAGTGTGTG	2002
37	Prior Application Number: 60/092182	QY	1441	GAGGACGCCACGCGGGAAGCTGAAGCTGCATAAGTGCAAGGGGCCCCCATGCGGCTGGCGGC	1500

Db 2003 GAGACCCCAACGGGGAAGCTGAAGCTGCATAAGTGAAGGCCCCCATGCGCTGGGGCGC 2062
Qy 1501 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTACGGGAGGCGAGGAGGCTGCACC 1560
Db 2063 AGCAGAGCCCTCTCCAACTCGTCCCAAGTACTACGGGAGGCGAGGAGGCTGCACC 2122
Qy 1561 TGTGACAGCGGGGACTACAAGCTCAGCTCGGCGGACGCGCGGAAAAAATCTTTCAAGAAG 1620
Db 2123 TGTGACAGCGGGGACTACAAGCTCAGCTCGGCGGACGCGCGGAAAAAATCTTTCAAGAAG 2182
Qy 1621 AAGTACAAGGCGAGCTATGTCGCGAGTCGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 1680
Db 2183 AAGTACAAGGCGAGCTATGTCGCGAGTCGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 2242
Qy 1681 GGCAGGGTGTACCAAGTGTAGGCTGGGTGATGCGCGCCAGCGCGGAAAACTTCACCAAGCGG 1740
Db 2243 GGCAGGGTGTACCAAGTGTAGGCTGGGTGATGCGCGCCAGCGCGGAAAACTTCACCAAGCGG 2302
Qy 1741 CACTGGCCAGGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 1800
Db 2303 CACTGGCCAGGGCCCTGAGGACCAAGATGACAGGATGGTGGGACTTCAGTGGCACT 2362
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCGCCAACTTAAAGTGACACATCGGTGTACATC 1860
Db 2363 GGAGGCTTCCCGACTACTCAGCGCGCCAACTTAAAGTGACACATCGGTGTACATC 2422
Qy 1861 CTAGAGAACACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCGAGGCTGGAAA 1920
Db 2423 CTAGAGAACACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCGAGGCTGGAAA 2482
Qy 1921 GACCACAAGCTGCACATCGACACAGAGATGAACCTCGCAGAACAAATTAAGAACCTG 1980
Db 2483 GACCACAAGCTGCACATCGACACAGAGATGAACCTCGCAGAACAAATTAAGAACCTG 2542
Qy 1981 AGGAACTCCGAGTGCACCTGAAAGAAAGCGCCAGAGAAATGTGACTGTCACAAAATC 2040
Db 2543 AGGAACTCCGAGTGCACCTGAAAGAAAGCGCCAGAGAAATGTGACTGTCACAAAATC 2602
Qy 2041 AGTACACACACGACGACCAAGGCGGCTCAAGCAGAGGCTCAGTGTGATCCTTTC 2100
Db 2603 AGTACACACACGACGACCAAGGCGGCTCAAGCAGAGGCTCAGTGTGATCCTTTC 2662
Qy 2101 AGGAAGGCTGCAAGAGAGAGCAAGGCTGTGCTGTGCGGAGCAGAGCGCAAGAG 2160
Db 2663 AGGAAGGCTGCAAGAGAGAGCAAGGCTGTGCTGTGCGGAGCAGAGCGCAAGAG 2722
Qy 2161 AACTCCGAGGCTGTCAAGCGCTGAGAACACAGACAGTGTGAGGATGCGAGGCTC 2220
Db 2723 AACTCCGAGGCTGTCAAGCGCTGAGAACACAGACAGTGTGAGGATGCGAGGCTC 2782
Qy 2221 ACGTGTTCACCCAGACACACAGCACTGGCAGACGCGGCTTTCTGGACACTGGGGCT 2280
Db 2783 ACGTGTTCACCCAGACACACAGCACTGGCAGACGCGGCTTTCTGGACACTGGGGCT 2842
Qy 2281 TTCTGTGCTGCAACGAGCGCAACAATAACAGTACTGTGTGATGAGGACCAATGAG 2340
Db 2843 TTCTGTGCTGCAACGAGCGCAACAATAACAGTACTGTGTGATGAGGACCAATGAG 2902
Qy 2341 ACTCACAATTTCTCTCTGTGATTTGCAATGCACTGGCTTCTAGAGTACTTGTATCTCAAC 2400
Db 2903 ACTCACAATTTCTCTCTGTGATTTGCAATGCACTGGCTTCTAGAGTACTTGTATCTCAAC 2962
Qy 2401 ACAGACCCCTACAGCTGTGATGAGTGAACACACACTGGACAGGATGTCTCAACACAG 2460
Db 2963 ACAGACCCCTACAGCTGTGATGAGTGAACACACACTGGACAGGATGTCTCAACACAG 3022
Qy 2461 CTACAGTACAGTCTAGGAGCTGAGAGCTGAAGGTTAAGAGTGTAAACCCCGG 2520
Db 3023 CTACAGTACAGTCTAGGAGCTGAGAGCTGAAGGTTAAGAGTGTAAACCCCGG 3082
Qy 2521 ACTCGAAACATGACCTGGGACTTAAAGATGGAGGAGCTATGAGCAATACAGGCACTTT 2580

Db 3083 ACTCGAAACATGAGACCTG-----CATGGAGGAAGCTATGAGCAATACAGGCACTTT 3133
Qy 2581 CAGCGTCCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
Db 3134 CAGCGTCCGAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193
Qy 2641 GAAGCTCGGAAGGTTAAGAAACCAACAGAGGTGACCTTCCAAAAACATAGAGGCACTCAC 2700
Db 3194 GAAGCTCGGAAGGTTAAGAAACCAACAGAGGTGACCTTCCAAAAACATAGAGGCACTCAC 3253
Qy 2701 TCATGTCACAGGCAATGAAGAAACCATGTGGTGTATTTCCAGCAGACCTGTCTATTGGCC 2760
Db 3254 TCATGTCACAGGCAATGAAGAAACCATGTGGTGTATTTCCAGCAGACCTGTCTATTGGCC 3313
Qy 2761 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3314 AGGAGGCTTGAGAAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
Qy 2821 CAGCAGGACGAGATTAACCTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 2880
Db 3374 CAGCAGGACGAGATTAACCTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGTTGGATTA 3433
Qy 2881 TACCTCACCGCTGCACAAAATGCAATTTTCTGATCAAAAAGTCAACACTTAACCTTCCC 2940
Db 3434 TACCTCACCGCTGCACAAAATGCAATTTTCTGATCAAAAAGTCAACACTTAACCTTCCC 3493
Qy 2941 CCAGAGCTCAAAAGGAAACCGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCAAAAGGAAACCGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553
Qy 3001 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3060
Db 3554 CCCAAGGCGGAAAGTCAATTTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3613
Qy 3061 TCCTCTTATTTCTTTTGGTTTGTCAAAAAGAGGAATTAAGAAGCAGGACAGAGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTTGGTTTGTCAAAAAGAGGAATTAAGAAGCAGGACAGAGCAACGT 3673
Qy 3121 GGAGAGCTGAAACAGTGCAGAGACGTTTGAATGACATAGTCAAGAGCAGGACAGAGATGA 3180
Db 3674 GGAGAGCTGAAACAGTGCAGAGACGTTTGAATGACATAGTCAAGAGCAGGACAGAGATGA 3733
Qy 3181 CATTTACTAGCATATAAACCCCTTGGTGTGCTCTGAGAAACTGCTTTCATTGTATATAT 3240
Db 3734 CATTTACTAGCATATAAACCCCTTGGTGTGCTCTGAGAAACTGCTTTCATTGTATATAT 3793
Qy 3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTCTAATAAGAAATCCCAA 3300
Db 3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAACTCTAATAAGAAATCCCAA 3853
Qy 3301 TTTTCAGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAAGAAAAA 3353
Db 3854 TTTTCAGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAAGAAAAA 3906

RESULT 14

US-09-991-073-83

; Sequence 83, Application US/09991073

; Patent No. US20202127576A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Daniel
APPLICANT: Tamas, Colin K.
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C15
CURRENT APPLICATION NUMBER: US/09/991,073
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC8
CURRENT APPLICATION NUMBER: US/09/990,442
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25

		PRIOR APPLICATION NUMBER: 60/090578	
		PRIOR FILING DATE: 1998-06-25	
		PRIOR APPLICATION NUMBER: 60/090690	
		PRIOR FILING DATE: 1998-06-25	
		PRIOR APPLICATION NUMBER: 60/090694	
		PRIOR FILING DATE: 1998-06-25	
		PRIOR APPLICATION NUMBER: 60/090695	
		PRIOR FILING DATE: 1998-06-25	
		PRIOR APPLICATION NUMBER: 60/090696	
		PRIOR FILING DATE: 1998-06-25	
		PRIOR APPLICATION NUMBER: 60/090862	
		PRIOR FILING DATE: 1998-06-26	
		PRIOR APPLICATION NUMBER: 60/090863	
		PRIOR FILING DATE: 1998-06-26	
		PRIOR APPLICATION NUMBER: 60/091360	
		PRIOR FILING DATE: 1998-07-01	
		PRIOR APPLICATION NUMBER: 60/091478	
		PRIOR FILING DATE: 1998-07-02	
		PRIOR APPLICATION NUMBER: 60/091544	
		PRIOR FILING DATE: 1998-07-01	
		PRIOR APPLICATION NUMBER: 60/091519	
		PRIOR FILING DATE: 1998-07-02	
		PRIOR APPLICATION NUMBER: 60/091626	
		PRIOR FILING DATE: 1998-07-02	
		PRIOR APPLICATION NUMBER: 60/091633	
		PRIOR FILING DATE: 1998-07-02	
		PRIOR APPLICATION NUMBER: 60/091978	
		PRIOR FILING DATE: 1998-07-07	
		PRIOR APPLICATION NUMBER: 60/091982	
		PRIOR FILING DATE: 1998-07-07	
		PRIOR APPLICATION NUMBER: 60/092182	
		PRIOR FILING DATE: 1998-07-09	
		Query Match 98.8%; Score 3323.4; DB 10; Length 3906;	
		Best Local Similarity 99.7%; Pred. No. 0;	
		Matches 3343; Conservative 1; Indels 9; Gaps 1;	
QY	1	GACTCCCGCATCCCAAGAACACACAGATCAGCAAAAAGAGATGGCCCGGAGC	60
DB	563	GACTCCCGCATCCCAAGAACACACAGATCAGCAAAAAGAGATGGCCCGGAGC	622
QY	61	CTCGTCTGTGCTGTGTCGCAACTGTGTTCTCCCTGCTGGTGGAGTTCGGCTTC	120
DB	623	CTCGTCTGTGCTGTGTCGCAACTGTGTTCTCCCTGCTGGTGGAGTTCGGCTTC	682
QY	121	CTGTCCACACCGCTGAAAGGACGTTTCAGAGGACCGCAGGAACATCGCCCAAC	180
DB	683	CTGTCCACACCGCTGAAAGGACGTTTCAGAGGACCGCAGGAACATCGCCCAAC	742
QY	181	ATCATCTCTGTGTGACGAGACACAGGATGTGGAGCTGGTTCATGCAAGTATGAAC	240
DB	743	ATCATCTCTGTGTGACGAGACACAGGATGTGGAGCTGGTTCATGCAAGTATGAAC	802
QY	241	AAGACCGGGGANTCATGAGCAGGGCGGGGCGACTTCATCAAGCCCTTGTGACCA	300
DB	803	AAGACCGGGGANTCATGAGCAGGGCGGGGCGACTTCATCAAGCCCTTGTGACCA	862
QY	301	CCCATGTGTCGCCCTCAGCTCCTCCATCTCTCACTGGCAAGTACGTCCACAACCAAC	360
DB	863	CCCATGTGTCGCCCTCAGCTCCTCCATCTCTCACTGGCAAGTACGTCCACAACCAAC	922
QY	361	ACCTACCAACAATGAGAACTGCTCTCCCTCTCGCCCTCTGTGACGCAAGCAGCGC	420
DB	923	ACCTACCAACAATGAGAACTGCTCTCCCTCTCGCCCTCTGTGACGCAAGCAGCGC	982
QY	421	ACCTTTGCGGTGTACCTCAATAGCACTGCTACCGGACAGCTTCTTCCGGGAAGTATCTT	480
DB	983	ACCTTTGCGGTGTACCTCAATAGCACTGCTACCGGACAGCTTCTTCCGGGAAGTATCTT	1042
QY	481	AATGAATACAGCGCTCTTACGTGCGCAACCGCGCTGGAAGAGTGGGTCCGACTCCTTAA	540
DB	1043	AATGAATACAGCGCTCTTACGTGCGCAACCGCGCTGGAAGAGTGGGTCCGACTCCTTAA	1102

QY	541	AACCTCCCGCTTTTATAACTACACGCTGTGTGGAACGGGGTGAAGAAAAGACCGCTTC	600
DB	1103	AACCTCCCGCTTTTATAACTACACGCTGTGTGGAACGGGGTGAAGAAAAGACCGCTTC	1162
QY	601	GACTACTCCAAGGATTACCTCACAGACCTCATCAATATGACAGGCTGAGCTTCTTCGC	660
DB	1163	GACTACTCCAAGGATTACCTCACAGACCTCATCAATATGACAGGCTGAGCTTCTTCGC	1222
QY	661	AGCTCCAAGAAGATGTACCCGACAGGCGAGCTCTCATGGTTCATCAGCCATCAGCCCC	720
DB	1223	AGCTCCAAGAAGATGTACCCGACAGGCGAGCTCTCATGGTTCATCAGCCATCAGCCCC	1282
QY	721	CACGGCCCTGAGATTACAGCCCAACATATTCACGGCTCTTCCCAACGCACTTCAGAC	780
DB	1283	CACGGCCCTGAGATTACAGCCCAACATATTCACGGCTCTTCCCAACGCACTTCAGAC	1342
QY	781	ATCACGCCGAGCTTACAACTACCGGCCCAACCCGGAACAAACACTGGATCATGCGCTACAG	840
DB	1343	ATCACGCCGAGCTTACAACTACCGGCCCAACCCGGAACAAACACTGGATCATGCGCTACAG	1402
QY	841	GGCCCATGAAGCCCATCCATGGAATTCACAACTGCTCCAGCGGAAGCGCTTCAG	900
DB	1403	GGCCCATGAAGCCCATCCATGGAATTCACAACTGCTCCAGCGGAAGCGCTTCAG	1462
QY	901	ACCTCATGTTCGCTGAGACGACTCCATGGAGACGATTTACAACTGCTGTTGAGCGGC	960
DB	1463	ACCTCATGTTCGCTGAGACGACTCCATGGAGACGATTTACAACTGCTGTTGAGCGGC	1522
QY	961	GAGCTGGACACACGCTTACATCTATATACACCGCGCACCGGTTACCACTCGCCAGTTT	1020
DB	1523	GAGCTGGACACACGCTTACATCTATATACACCGCGCACCGGTTACCACTCGCCAGTTT	1582
QY	1021	GGCTGCTGAAAGGAAATCCATGCTATGATTTGACATCAGGCTCCGTTCTACGCTG	1080
DB	1583	GGCTGCTGAAAGGAAATCCATGCTATGATTTGACATCAGGCTCCGTTCTACGCTG	1642
QY	1081	AGGGGCCCCAAAGTGGAAGCGGCTGTCTGTAATCCCCACATCGTCTCTCAACTGACCTG	1140
DB	1643	AGGGGCCCCAAAGTGGAAGCGGCTGTCTGTAATCCCCACATCGTCTCTCAACTGACCTG	1702
QY	1141	GCCTCCACATCTGACATATTCAGGCTTGGACATACCTGCGGATATGACCGGGAATCC	1200
DB	1703	GCCTCCACATCTGACATATTCAGGCTTGGACATACCTGCGGATATGACCGGGAATCC	1762
QY	1201	ATCTCAAGCTGTGACACGAGCGCGGCTGTAATCGGTTTCACTTGAAGAAAGATG	1260
DB	1763	ATCTCAAGCTGTGACACGAGCGCGGCTGTAATCGGTTTCACTTGAAGAAAGATG	1822
QY	1261	AGGCTCTGCGGAGCTCTCTTCTGTTGGAGAGAGGCAAGCTGCTTACAAAGAGACAAT	1320
DB	1823	AGGCTCTGCGGAGCTCTCTTCTGTTGGAGAGAGGCAAGCTGCTTACAAAGAGACAAT	1882
QY	1321	GACAAGTGGACGCCCGAGGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGGACCTG	1380
DB	1883	GACAAGTGGACGCCCGAGGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGGACCTG	1942
QY	1381	TGTGAGCTGTGAGTACCAAGCGGCTGTGACAGCTGGGACAGAGTGGAGTGTGTG	1440
DB	1943	TGTGAGCTGTGAGTACCAAGCGGCTGTGACAGCTGGGACAGAGTGGAGTGTGTG	2002
QY	1441	GAGGACGCCACCGGGAGCTGAGCTGCATAGTGCAAGGGCCCACTGCGGCTGGCGGC	1500
DB	2003	GAGGACGCCACCGGGAGCTGAGCTGCATAGTGCAAGGGCCCACTGCGGCTGGCGGC	2062
QY	1501	AGCAGAGCCCTCTCCAACTCTGTCGCCAAAGTACTACGGGACGGGCAAGCGGCTTSCAC	1560
DB	2063	AGCAGAGCCCTCTCCAACTCTGTCGCCAAAGTACTACGGGACGGGCAAGCGGCTTSCAC	2122
QY	1561	TGTGACAGCGGGACCTCAAGCTCAGCTGGCGGACCGGGGAAAACCTCTTCAGAGAG	1620
DB	2123	TGTGACAGCGGGACCTCAAGCTCAGCTGGCGGACCGGGGAAAACCTCTTCAGAGAG	2182
QY	1621	AAGTACAGGCCAGCTATGTCCGAGTCTCCATCCGCTCAGTGGCCATCGAGGTGGAC	1680

Db 2183 AGTACAAAGGCGGAGCTATGTCGCGAGTCGCTCCATCGCTCAGTGGCCATCGAGGTGGAC 2242
Qy 1681 GGCAGGGGTGTACCAAGTGTAGGCTGGGTGATGCGCGCCAGCCCGGAAACCTTCACCAAGCGG 1740
Db 2243 GGCAGGGGTGTACCAAGTGTAGGCTGGGTGATGCGCGCCAGCCCGGAAACCTTCACCAAGCGG 2302
Qy 1741 CACTGGCCAGGGGCCCTGTAGGACCAAGATGACAAGATGGTGGGACTTCAGTGGCACT 1800
Db 2303 CACTGGCCAGGGGCCCTGTAGGACCAAGATGACAAGATGGTGGGACTTCAGTGGCACT 2362
Qy 1801 GGAGGCGCTTCCGAGCTACTCAGCGCGCCCAACCCATTAAGTGTACATCGGTGCTACATC 1860
Db 2363 GGAGGCGCTTCCGAGCTACTCAGCGCGCCCAACCCATTAAGTGTACATCGGTGCTACATC 2422
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTGCAGGCGCTGGAAA 1920
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTGCAGGCGCTGGAAA 2482
Qy 1921 GACCACAGCTGCACATCGACCCAGAGATTGAACCTTGCAGAACAAATTAAGAACCTG 1980
Db 2483 GACCACAGCTGCACATCGACCCAGAGATTGAACCTTGCAGAACAAATTAAGAACCTG 2542
Qy 1981 AGGGAAGTCCGAGGTCACTGAAAGAAAAGCGGCCAGAGATGTGACTGTCACAAATC 2040
Db 2543 AGGGAAGTCCGAGGTCACTGAAAGAAAAGCGGCCAGAGATGTGACTGTCACAAATC 2602
Qy 2041 AGTACACACCCAGCAAAAGCGCGCTCAAGCAGAGAGCTTCCAGTGTGATCCTTTTC 2100
Db 2603 AGTACACACCCAGCAAAAGCGCGCTCAAGCAGAGAGCTTCCAGTGTGATCCTTTTC 2662
Qy 2101 AGGAAGGGCTGCAAGAGAGAGCAAGTGTGCTGTGCGGAGCAGAGCGCAGAGAG 2160
Db 2663 AGGAAGGGCTGCAAGAGAGAGCAAGTGTGCTGTGCGGAGCAGAGCGCAGAGAG 2722
Qy 2161 AAATCCGCAAGTGTCTCAAGCGCTTGCAGAACACAGACAGTGCAGAGTGCAGGCGCTC 2220
Db 2723 AAATCCGCAAGTGTCTCAAGCGCTTGCAGAACACAGACAGTGCAGAGTGCAGGCGCTC 2782
Qy 2221 ACGTGCTTACCCAGCACACAGCACTGCGACAGCGCGCTTCTTGGACACTGGGGCT 2280
Db 2783 ACGTGCTTACCCAGCACACAGCACTGCGACAGCGCGCTTCTTGGACACTGGGGCT 2842
Qy 2281 TTCTGTGCTGCGACAGCGCAACAATAACACGCTACTGTGTCATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGCGACAGCGCAACAATAACACGCTACTGTGTCATGAGGACCATCAATGAG 2902
Qy 2341 ACTCAAAATTTCTCTCTGTGAATTTGCAACTGGGCTTCTAGAGTACTTTGATCTCAAC 2400
Db 2903 ACTCAAAATTTCTCTCTGTGAATTTGCAACTGGGCTTCTAGAGTACTTTGATCTCAAC 2962
Qy 2401 ACAGACCCCTACAGCTGTGATGATGAGTGAACACACACTGGACAGGATGCTCTCAACCG 2460
Db 2963 ACAGACCCCTACAGCTGTGATGATGAGTGAACACACACTGGACAGGATGCTCTCAACCG 3022
Qy 2461 CTACAGCTACAGTCTATGGAGCTGAGAGCTGCAAGGGTTTACAAGCAGTGTAAACCCCGG 2520
Db 3023 CTACAGCTACAGTCTATGGAGCTGAGAGCTGCAAGGGTTTACAAGCAGTGTAAACCCCGG 3082
Qy 2521 ACTCGAACAATGGACCTGGGACTTAAAGTGGAGGAGCTATGAGCAATACAGGCAATTT 2580
Db 3083 ACTCGAACAATGGACCTGGG-----GATGGGAGGAGCTATGAGCAATACAGGCAATTT 3133
Qy 2581 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCAAATCACTGGGACAACTGTGG 2640
Db 3134 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTTCTTCAAATCACTGGGACAACTGTGG 3193
Qy 2641 GRAGGCTGGGAAGGTTAAGAAAACAAGAGGTGGAACCTTCCAAAAAATAGAGGCAATCAC 2700
Db 3194 GAAGGCTGGGAAGGTTAAGAAAACAAGAGGTGGAACCTTCCAAAAAATAGAGGCAATCAC 3253
Qy 2701 TGACTGCACAGGCAATGAAAAACCATGTGGTGCATTTTCCAGCAGACCTGTGTATTGGCC 2760

Db 3254 TGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTCTATTGGCC 3313
Qy 2761 AGGAGGCTGTGAAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 2820
Db 3314 AGGAGGCTGTGAAAAGCAAGCAAGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAAC 3373
Qy 2821 CAGCAGGACGAGAGATAACTTCAGGAAGTCAATTTTGGCCCTGCTTTTGGTATTGGAATTA 2880
Db 3374 CAGCAGGACGAGAGATAACTTCAGGAAGTCAATTTTGGCCCTGCTTTTGGTATTGGAATTA 3433
Qy 2881 TACTCTACCGAGCTGCACAAATGCAATTTTCTGATCAAAAGTCAACATTAACCCCTCCC 2940
Db 3434 TACTCTACCGAGCTGCACAAATGCAATTTTCTGATCAAAAGTCAACATTAACCCCTCCC 3493
Qy 2941 CCAGAGCTCAAAAGGAAAACCGAGAGAGCGAGAGAGATTTCTCTGGAAATTTCT 3000
Db 3494 CCAGAGCTCAAAAGGAAAACCGAGAGAGCGAGAGAGATTTCTCTGGAAATTTCT 3553
Qy 3001 CCCAAGGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGAAAAGCAGTCTGTCTTAAA 3060
Db 3554 CCCAAGGGCGAAAGTCAATTTGGAATTTTAAATCATAGGGAAAAGCAGTCTGTCTTAAA 3613
Qy 3061 TCCTCTTATTCTTTTGGTTTGTCAAAAGAGGAACCTAAAGAGCAGCAGAGCAACGT 3120
Db 3614 TCCTCTTATTCTTTTGGTTTGTCAAAAGAGGAACCTAAAGAGCAGCAGAGCAACGT 3673
Qy 3121 GGAGAGGCTGAAAACAGTGCAGAGACCTTTTGAACAATGAGTCAAGTCAAGAGAGATGA 3180
Db 3674 GGAGAGGCTGAAAACAGTGCAGAGACCTTTTGAACAATGAGTCAAGTCAAGAGAGATGA 3733
Qy 3181 CATTTTACCTAGCATAAAACCTTGGCTCTGAGAGAACTGCCCTTCATTTGATATAT 3240
Db 3734 CATTTTACCTAGCATAAAACCTTGGCTCTGAGAGAACTGCCCTTCATTTGATATAT 3793
Qy 3241 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAAACCTTAATAAGAAATCCCAA 3300
Db 3794 GTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGAAACCTTAATAAGAAATCCCAA 3853
Qy 3301 TTTTTCAGAGTGGTGGTGTCAATTAACCGCTCTGTGGCCAGTGTAAAGAAAAA 3353
Db 3854 TTTTTCAGAGTGGTGGTGTCAATTAACCGCTCTGTGGCCAGTGTAAAGAAAAA 3906

Search completed: February 14, 2004, 22:52:51
Job time : 1148 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 11:41:13 ; Search time 11897 Seconds
(without alignments)
11564.182 Million cell updates/sec

Title: US-10-025-966A-4
Perfect score: 3363
Sequence: 1 gactccgcatcccaaga.....aaagaaaaaa 3363

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ses.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_ses.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_man.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3363	100.0	3363	6	AX573182	AX573182 Sequence
2	3347.2	99.5	4397	9	AB033073	AB033073 Homo sapi
3	3346.8	99.5	4279	9	AY101176	AY101176 Homo sapi
4	3346.8	99.5	4286	6	AX573191	AX573191 Sequence
5	3344.8	99.5	3348	6	AX211371	AX211371 Sequence
6	3323.4	98.8	3906	6	AR252433	AR252433 Sequence
7	3323.4	98.8	3906	6	AX403196	AX403196 Sequence
8	3323.4	98.8	3906	6	AX454488	AX454488 Sequence
9	3323.4	98.8	3906	6	AX463886	AX463886 Sequence
10	3323.4	98.8	3906	6	AX490966	AX490966 Sequence
11	3223.2	95.8	3513	6	AX470095	AX470095 Sequence
12	2654.8	78.9	3289	6	AX573192	AX573192 Sequence
13	2613	77.7	2616	6	AX573183	AX573183 Sequence
14	2605.2	77.5	2610	6	AX470097	AX470097 Sequence
15	2294.8	68.2	2559	6	BD127910	BD127910 Primer fo
16	2294.8	68.2	2559	9	AK074483	AK074483 Homo sapi
17	2174.8	64.7	2178	6	BD127849	BD127849 Primer fo
18	2174.8	64.7	2178	9	AK075035	AK075035 Homo sapi
19	2095.2	62.3	3603	10	AY101177	AY101177 Mus muscu
20	2095.2	62.3	3613	6	AX573194	AX573194 Sequence
21	2095.2	62.3	3766	6	AX573188	AX573188 Sequence
22	2058.2	61.2	2628	6	AX573189	AX573189 Sequence
23	2058.2	61.2	2628	6	AX573195	AX573195 Sequence
24	1727.4	51.4	1729	9	HS049G161	AL133001 Novel hum
25	1625	48.3	1661	9	BC020962	BC020962 Homo sapi
26	1151	34.2	3688	5	BC045403	BC045403 Danio rer
27	1080.4	32.1	4088	10	AF230072	AF230072 Rattus no
28	1070.6	31.8	5769	5	AF410802	AF410802 Coturnix
29	1062.4	31.6	4623	10	BC034547	BC034547 Mus muscu
30	1062.4	31.6	4623	10	BC049376	BC049376 Mus muscu
31	1062.4	31.6	4671	10	AY101178	AY101178 Mus muscu
32	1040.4	30.9	2616	6	AX206971	AX206971 Sequence
33	1040.4	30.9	2616	6	AX573180	AX573180 Sequence
34	1040.4	30.9	4321	6	AX206963	AX206963 Sequence
35	1040.4	30.9	4378	6	AX188060	AX188060 Sequence
36	1040.4	30.9	4378	6	AX188065	AX188065 Sequence
37	1040.4	30.9	4378	6	AX188144	AX188144 Sequence
38	1040.4	30.9	4378	6	AX188145	AX188145 Sequence
39	1040.4	30.9	4378	6	AX188147	AX188147 Sequence
40	1040.4	30.9	4378	6	AX188329	AX188329 Sequence
41	1040.4	30.9	4378	6	AX188343	AX188343 Sequence
42	1040.4	30.9	4669	9	AY101175	AY101175 Homo sapi
43	1040.4	30.9	4961	6	AX463510	AX463510 Sequence
44	1040.4	30.9	5244	6	AX573179	AX573179 Sequence
45	1038.8	30.9	5699	9	AF545571	AF545571 Homo sapi

ALIGNMENTS

RESULT 1	AX573182	AX573182	3363 bp	DNA	linear	PAT 29-NOV-2002
LOCUS	AX573182	Sequence 4 from Patent WO02059327.				
DEFINITION	AX573182	Sequence 4 from Patent WO02059327.				
ACCESSION	AX573182	GI:26005103				
VERSION	AX573182.1					
KEYWORDS		Homo sapiens (human)				
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE 1						
AUTHORS		Rosen,S.D., Hemmerich,S., Tomita,M. and Palmeri,D.				
TITLE		Sulfatases and methods of use thereof				
JOURNAL		Patent: WO 02059327-A 4 01-AUG-2002;				

1981 AGGGAAGTCCAGAGTCACTGTAAGAAAGGCGCCAGAAAGATGTGACTGTGACAAAATC 2040
2041 AGCTACACACACCCAGACAAAAGGCGCCTCAAGCAGAGAGGTCCAGTCTGCAATCTTTC 2100
2041 AGCTACACACACCCAGACAAAAGGCGCCTCAAGCAGAGAGGTCCAGTCTGCAATCTTTC 2100
2101 AGGAAGGCGCTGCAAGAGAGGAGCAAGAGTGTGGCTTTGCGGAGCAGAGGCGCAAGAG 2160
2101 AGGAAGGCGCTGCAAGAGAGGAGCAAGAGTGTGGCTTTGCGGAGCAGAGGCGCAAGAG 2160
2161 AAATCGCGAAGTGTCTCAAGCGCCTGCGAGCAACAGCAGCAGTGCAGCATGCCAGGCCTC 2220
2161 AAATCGCGAAGTGTCTCAAGCGCCTGCGAGCAACAGCAGCAGTGCAGCATGCCAGGCCTC 2220
2221 ACCTGCTTCAACCCAGACACACAGCACTGGCAGACGGCGCTTTCTGGGACACATGGGGCT 2280
2281 TTCTGTGCTGCGACCGCCACCAATACACGTAAGTGTGCTGATGAGGACCATCATGAG 2340
2281 TTCTGTGCTGCGACCGCCACCAATACACGTAAGTGTGCTGATGAGGACCATCATGAG 2340
2341 ACTCACAATTTCTCTCTGTAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2400
2341 ACTCACAATTTCTCTCTGTAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAAC 2400
2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGCAAGAGGATGCTCTCAACGAG 2460
2401 ACAGACCCCTACAGCTGATGAATGCAAGTGAACACACTGCAAGAGGATGCTCTCAACGAG 2460
2461 CTACAGTCAAGTCTCATGAGAGTGAAGGCTGCAAGGTTTACAAGCAGTGAACCCCGG 2520
2461 CTACAGTCAAGTCTCATGAGAGTGAAGGCTGCAAGGTTTACAAGCAGTGAACCCCGG 2520
2521 ACTCGAAATGACCTGGACCTTAAAGTGGAGGAGCTATGAGCAATACAGGCACTTT 2580
2521 ACTCGAAATGACCTGGACCTTAAAGTGGAGGAGCTATGAGCAATACAGGCACTTT 2580
2581 CAGGTCGAAGTGGCAGAAATGAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
2581 CAGGTCGAAGTGGCAGAAATGAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
2641 GAAGGCTGGGAAGTTAAGAAACACAGAGTGGAGCTCCAAACATAGAGGATCAAC 2700
2641 GAAGGCTGGGAAGTTAAGAAACACAGAGTGGAGCTCCAAACATAGAGGATCAAC 2700
2701 TGACTGACAGGCAATGAAACACATGTGGTGAATTCAGCAGACCTGTGCTATGGCC 2760
2701 TGACTGACAGGCAATGAAACACATGTGGTGAATTCAGCAGACCTGTGCTATGGCC 2760
2761 AGGAGGCTCAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTCCTGGAGGATAAC 2820
2761 AGGAGGCTCAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTCCTGGAGGATAAC 2820
2821 CAGCAGGAGCAGAGATAAATTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTGGATTA 2880
2821 CAGCAGGAGCAGAGATAAATTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTGGATTA 2880
2881 TACTCAGCAGCTGCACAAATGCAATTTTTCATCAAAAGTCCACCACTCAACCTCC 2940
2881 TACTCAGCAGCTGCACAAATGCAATTTTTCATCAAAAGTCCACCACTCAACCTCC 2940
2941 CCAGAGCTCACAAGGAAACGAGAGCGGAGAGAGAGATTTTCCCTTGGAAATTTCT 3000
2941 CCAGAGCTCACAAGGAAACGAGAGCGGAGAGAGAGATTTTCCCTTGGAAATTTCT 3000
3001 CCAAGGCGAAGTCAATGGAATTTTAAATCATAGGGAAGAGCAGTCTGCTTTCTAA 3060
3001 CCAAGGCGAAGTCAATGGAATTTTAAATCATAGGGAAGAGCAGTCTGCTTTCTAA 3060
3061 TCCTCTTATTTCTTTGTTGTCAAAAGAGGAACCTAAGAACGAGCAGACAGAGCAACGT 3120

Db 3061 TCCTCTTATTTCTTTTGTGTGTCACAAAGAGGAACCTAAGAGCAGGACAGAGCAACGT 3120
QY 3121 GGAGAGCTGAAACACAGTCAGAGAGACGTTTTCACAAATGAGTCAGTCAGCAAAAGAGATGA 3180
Db 3121 GGAGAGCTGAAACACAGTCAGAGAGACGTTTTCACAAATGAGTCAGTCAGCAAAAGAGATGA 3180
QY 3181 CATTACTAGCACTATAAACCCCTGGTTCCTCTGAAGAACTGCCTTCATTGCTATATAT 3240
Db 3181 CATTACTAGCACTATAAACCCCTGGTTCCTCTGAAGAACTGCCTTCATTGCTATATAT 3240
QY 3241 GTGACTATTACATGTAATCAACATGGAACCTTTTAGGGGAACCTAATAAGAAATCCCAA 3300
Db 3241 GTGACTATTACATGTAATCAACATGGAACCTTTTAGGGGAACCTAATAAGAAATCCCAA 3300
QY 3301 TTTTCAGAGTGGTGTGTCATAAACCGCTCTGTGGCCAGTGTAAAGAAAGAAAAA 3360
Db 3301 TTTTCAGAGTGGTGTGTCATAAACCGCTCTGTGGCCAGTGTAAAGAAAGAAAAA 3360
QY 3361 AAA 3363
Db 3361 AAA 3363

RESULT 2
AB033073
LOCUS Homo sapiens mRNA for KIAA1247 protein, partial cds.
DEFINITION Homo sapiens mRNA for KIAA1247 protein, partial cds.
ACCESSION AB033073
VERSION AB033073.2 GI:14133244
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
Nagase,T., Ishikawa,K., Kikuno,R., Hirotsawa,M., Momura,N. and Ohara,O.
Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
DNA Res. 6 (5), 337-345 (1999)
JOURNAL MEDLINE 20039619
PUBMED 10574462
REFERENCE 2 (bases 1 to 4397)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (04-OCT-1999) Otsu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@kzusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
On May 17, 2001 this sequence version replaced gi:6330839.
COMMENT Location/Qualifiers
FEATURES
source
1..4397
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pJ01452"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
/note="vector:pBluescriptII SK plus. This sequence is
replaced that of h00128a cDNA as a representative cDNA
sequence for KIAA1247"
1..4397
/gene="KIAA1247"
/gene="KIAA1247"
/note="KIAA1247"
/note="KIAA1247 protein"
/product="KIAA1247 protein"
/protein_id="BA08561.2"
/db_xref="GI:14133245"
/translation="DSRIKPEAPDQKKWPPSLVLCILLSATVFLSGSSAFLSHH"

1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTGACAGGCTCGAAA 1920
2146 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAAGTCCCTGACAGGCTCGAAA 2205
1921 GACCACAAAGCTGCACATCGACACGAGATTCACACCTGACAGACACAAATTAAGAACCTG 1980
2206 GACCACAAAGCTGCACATCGACACGAGATTCACACCTGACAGACACAAATTAAGAACCTG 2265
1981 AGGAAAGTCCGAGTGCACCTGAAGAAAAGCGGCAGAGAATGTGACTGTCCAAAATC 2040
2266 AGGAAAGTCCGAGTGCACCTGAAGAAAAGCGGCAGAGAATGTGACTGTCCAAAATC 2325
2041 AGCTACACACCCGACACAAAGGCGGCTCAAGCAGAGAGCTCCAGTCTGCATCCTTTC 2100
2326 AGCTACACACCCGACACAAAGGCGGCTCAAGCAGAGAGCTCCAGTCTGCATCCTTTC 2385
2101 AGGAAGGCGCTGCAAGAGAGGACAGAGTGTGGCTGTGTGGGAGCAGAGAAGCGCAAGAAG 2160
2386 AGGAAGGCGCTGCAAGAGAGGACAGAGTGTGGCTGTGTGGGAGCAGAGAAGCGCAAGAAG 2445
2161 AAACCTCGCAAGCTGTCAAGCGCTCGAGAACAAACGACACGTCGACATGCCAGGCTC 2220
2446 AAACCTCGCAAGCTGTCAAGCGCTCGAGAACAAACGACACGTCGACATGCCAGGCTC 2505
2221 ACGTGCTTACCCACGACCAACGACACTGGCAGACGCGCTTCTGAGACACTGGGCGCT 2280
2506 ACGTGCTTACCCACGACCAACGACACTGGCAGACGCGCTTCTGAGACACTGGGCGCT 2565
2281 TTCTGTGCTGTCACAGCGCCCAACAAATAACAGTACTGTGTCATGAGGACCAATCAATGAG 2340
2566 TTCTGTGCTGTCACAGCGCCCAACAAATAACAGTACTGTGTCATGAGGACCAATCAATGAG 2625
2341 ACTCACAAATTCCTCTCTGTAATTTGCACTGGCTTCTAGAGTACTTTGATCTCAAC 2400
2626 ACTCACAAATTCCTCTCTGTAATTTGCACTGGCTTCTAGAGTACTTTGATCTCAAC 2685
2401 ACAGACCCCTACAGCTGATGAATCAGTGAACACACTGGACAGGAGTGTCTCTCAACAG 2460
2686 ACAGACCCCTACAGCTGATGAATCAGTGAACACACTGGACAGGAGTGTCTCTCAACAG 2745
2461 CTACACGTACAGCTCATGAGCTGAGAGCTGACAGGTTACAGAGTGTACACCCCGG 2520
2746 CTACACGTACAGCTCATGAGCTGAGAGCTGACAGGTTACAGAGTGTACACCCCGG 2805
2521 ACTCGAAACATGGAACCTGGGACTTAAAGATGAGGAAGCTATGACGAATACAGGCACTT 2580
2806 ACTCGAAACATGGAACCTGGGACTTAAAGATGAGGAAGCTATGACGAATACAGGCACTT 2865
2581 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCCTGGGACAACTGTGG 2640
2866 CAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCCTGGGACAACTGTGG 2925
2641 GAAGGCTGGGAAGGTTAAGAACACACAGAGTGGACCTCCAAACACATAGAGCATCACC 2700
2926 GAAGGCTGGGAAGGTTAAGAACACACAGAGTGGACCTCCAAACACATAGAGCATCACC 2985
2701 TGACTGCACAGGCAATGAAACACCATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
2986 TGACTGCACAGGCAATGAAACACCATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3045
2761 AGGAGGCTGAGAAAGCAGACGACCTCTCAGTCAACATGACAGATTCGGAGGATAAC 2820
3046 AGGAGGCTGAGAAAGCAGACGACCTCTCAGTCAACATGACAGATTCGGAGGATAAC 3105
2821 CAGCAGGACGACAGATACTTACAGAGTCCATTTTGGCCCTGCTTTGCTTGGATTA 2880
3106 CAGCAGGACGACAGATACTTACAGAGTCCATTTTGGCCCTGCTTTGCTTGGATTA 3165
2881 TACCTCAGCAGCTGCAAAAATGATTTTTCGTATCAAAAAGTCAACACTAACCCCTCCC 2940
3166 TACCTCAGCAGCTGCAAAAATGATTTTTCGTATCAAAAAGTCAACACTAACCCCTCCC 3225
2941 CCAGAGCTCACAAGGAAAACGGAGAGCGGAGAGAGATTTCTCTTGGAAATTTCT 3000

3226 CCAGAAGCTCACAAGGAAAACGGAGAGCGAGAGAGATTTCTTGAAAATTTCT 3285
3001 CCCAAGGCGCAAAAGTCATTGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTTCTAAA 3060
3286 CCCAAGGCGCAAAAGTCATTGGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTGTTCTAAA 3345
3061 TCTCTTATCTTTTGGTTTGTCAAAAGAGGAACTAAGAAAGCAGGACAGGCAACGT 3120
3346 TCTCTTATCTTTTGGTTTGTCAAAAGAGGAACTAAGAAAGCAGGACAGGCAACGT 3405
3121 GGAGAGGCTGAAAACAGTCAGAGACGTTTCACAATGAGTCAGTAGCACAAAAGAGATGA 3180
3406 GGAGAGGCTGAAAACAGTCAGAGACGTTTCACAATGAGTCAGTAGCACAAAAGAGATGA 3465
3181 CATTTACCTAGCATTAAACCTCGTTGCTCTCAAGAAAACCTGCTTCAATTGATATAT 3240
3466 CATTTACCTAGCATTAAACCTCGTTGCTCTCAAGAAAACCTGCTTCAATTGATATAT 3525
3241 GTGACTATTTACATGTATCAACATGGGAACTTTTAGGGGAACTTAAGAAAATCCCAA 3300
3526 GTGACTATTTACATGTATCAACATGGGAACTTTTAGGGGAACTTAAGAAAATCCCAA 3585
3301 TTTTCAGAGTGGTGTCTCAATAAACGCTCTGTGGCCAGTGTAAAGAAA 3352
3586 TTTTCAGAGTGGTGTCTCAATAAACGCTCTGTGGCCAGTGTAAAGAAA 3637

RESULT 3
AY101176
LOCUS
DEFINITION Homo sapiens extracellular sulfatase SULF-2 mRNA, complete cds.
ACCESSION AY101176
VERSION AY101176.1 GI:27356933
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Morimoto-Tomita,M., Uchimura,K., Werb,Z., Hemmerich,S. and Rosen,S.D.
TITLE Cloning and Characterization of Two Extracellular Heparin-degrading Endosulfatases in Mice and Humans
J. Biol. Chem. 277 (51), 49175-49185 (2002)
PUBMED 12368295
REFERENCE 2 (bases 1 to 4279)
AUTHORS Morimoto-Tomita,M., Werb,Z., Hemmerich,S. and Rosen,S.D.
TITLE Direct Submission
SUBMITTED (30-APR-2002) Biology, Thios Pharmaceuticals Inc., 747
Fifty Second Street, Oakland, CA 94609, USA
FEATURES
Location/Qualifiers
1. 4279
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.12"
/tissue type="lung"
670..3282
/codon_start=1
/product="extracellular sulfatase SULF-2"
/protein_id="AA076861.1"
/db_xref="GI:27356934"

CDS
1
/translat ion="MGPPSLVLCILSATVPSLGGSAFASHRLKGRFQDRNRNP
NIIIVTDQDELGVSNVKNRIMBOGGAHPINAFVTPMCCPSRSSILTKYVH
NNTVNNENSPSWOAHESRTFAVLNSTGYRTAFKYLNEYNGSVYPPQKWE
VGLKNSRPVNTLCRNGKEKESDYKDLTDLTNSVSFFRTSKMYPHPVLM
VISHAPGPDPSAFQSRLEFPNASQHTPSYINAPDKHWNRYTGPMPHPMEFT
NMLQRKRLTMSVDDSNETIYNMLVETGELDNTIYITADHGHYIQFGLVKGKSWP
YEFDIRFVYRGNFENCLPHIIVLNIDLAFTIADLIDIPADMGKSLKLLDT
ERPYNRFLKKKMRWRDSTFLVERKLIHKRDNDKVDQAQENFLPKYQVRKDLQRAE
YOTACEQLGQKWCQVEDATGKLKHKCKPMLGSRALSINLVKPYIQGSEACTCDS

GDYKLSLAGRRKKLKKYKASVRSRSIRSAIEVDGRVYHVLGDAAPRNLTKRH
WFGAEDQDDXGGSPSGTGGIPDYSAANPIKVTCHRYILENDTVQCDLILYKQAW
KDHKLHI DEIETLONKIKNLREVHGLKKRPEEDCHDKI SYTHKHGRKLKRGSSL
HPRKLGLOKDKVLLREOKRKKLLKLLRLONNDTCSMPGLTCTFDHNOHWOTAPF
WTLGPFCATCSANNITWCMFTINETHNLFCEPATGFELEVDLNTDPOYOLMNAWIL
DRDLNLQHLVQMLRSLCKGYKQCNPRTRNDMLGLKDGGSYEQYRQFORRWPKRKP
SSKSLGQIMEGWEG"
polyA signal
BASE COUNT 3944.3948
ORIGIN 1186 a 1143 c 1105 g 845 t

Query Match 99.5%; Score 3346.8; DB 9; Length 4279;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 GACTCCCGCATCCCAAAGAAGCACAGATCAGCAAAAAGAGATGGCCCCCGGAGC 60
DB 625 GACTCCCGCATCCCAAAGAAGCACAGATCAGCAAAAAGAGATGGCCCCCGGAGC 684
QY 61 CTCGTGCTGTGTGCTGTCCGCACTGTCTCCCTGCTGGTGGAGCTCGGCCTTC 120
DB 685 CTCGTGCTGTGTGCTGTCCGCACTGTCTCCCTGCTGGTGGAGCTCGGCCTTC 744
QY 121 CTGTCGACACACCGCTGAAAGCAGGTTTCAGAGGACCGCAGGAACATCGCCCCCAAC 180
DB 745 CTGTCGACACACCGCTGAAAGCAGGTTTCAGAGGACCGCAGGAACATCGCCCCCAAC 804
QY 181 ATCATCTGTGTGTCAGGACGACGACGAGTGTGAGCTGGTTCATGAGTGATGAAC 240
DB 805 ATCATCTGTGTGTCAGGACGACGACGAGTGTGAGCTGGTTCATGAGTGATGAAC 864
QY 241 AAGACCCGGCGCATCATGAGCAGGCGGGGGCGCACTTCATCAAGCCCTTCGTGACCA 300
DB 865 AAGACCCGGCGCATCATGAGCAGGCGGGGGCGCACTTCATCAAGCCCTTCGTGACCA 924
QY 301 CCATGTGTGCCCCCTCAAGCTCTCCATCTCACTGCAAGTACGTGCAAGTACGTGCA 360
DB 925 CCATGTGTGCCCCCTCAAGCTCTCCATCTCACTGCAAGTACGTGCAAGTACGTGCA 984
QY 361 ACCTACACCAAGTACAGACTCTCTCGCCCTCTGCGAGGACACAGAGAGCGCG 420
DB 985 ACCTACACCAAGTACAGACTCTCTCGCCCTCTGCGAGGACACAGAGAGCGCG 1044
QY 421 ACCTTTGCGGTGATCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCT 480
DB 1045 ACCTTTGCGGTGATCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAGTATCT 1104
QY 481 AATGAATACAGCGCTCTACGTGACACCGCGTGGAGAGTGGTGGACTCTTAA 540
DB 1105 AATGAATACAGCGCTCTACGTGACACCGCGTGGAGAGTGGTGGACTCTTAA 1164
QY 541 AACTCCCGCTTTTATACTACAGCTGTGTGGAACGGGGTGAAGAAAGCAGCGCTCC 600
DB 1165 AACTCCCGCTTTTATACTACAGCTGTGTGGAACGGGGTGAAGAAAGCAGCGCTCC 1224
QY 601 GATCTCCAGGATTAATCTCAGACCTCATCAACATGACAGCGTGGTTCCTCCG 660
DB 1225 GATCTCCAGGATTAATCTCAGACCTCATCAACATGACAGCGTGGTTCCTCCG 1284
QY 661 ACCTTCAAGAAGATGACCGGACAGCGCAGTCTCATGCTCATCAGCATGCGCCCC 720
DB 1285 ACCTTCAAGAAGATGACCGGACAGCGCAGTCTCATGCTCATCAGCATGCGCCCC 1344
QY 721 CACGGCCCTGAGGATTCAGCCCCACAAATATTCAGCCTCTTCCCAAACGCACTCAGCAC 780
DB 1345 CACGGCCCTGAGGATTCAGCCCCACAAATATTCAGCCTCTTCCCAAACGCACTCAGCAC 1404
QY 781 ATCAGCCGAGCTTAACTACGGCCGACCGGACCAACACTGGATCATGCGCTACAGC 840
DB 1405 ATCAGCCGAGCTTAACTACGGCCGACCGGACCAACACTGGATCATGCGCTACAGC 1464
QY 841 GGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGAG 900

1465 GGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGAAGCGCTTGCA 1524
QY 901 ACCTCATGTGCGGTGACACATCCATGAGAGAGATTTAAACATGCTGGTTGAGACGGCG 960
DB 1525 ACCTCATGTGCGGTGACACATCCATGAGAGAGATTTAAACATGCTGGTTGAGACGGCG 1584
QY 961 GAGCTGGAACACAGCTACATGCTATACACCGCGACACCGGTTACACATCGGCCAGTTT 1020
DB 1585 GAGCTGGAACACAGCTACATGCTATACACCGCGACACCGGTTACACATCGGCCAGTTT 1644
QY 1021 GGGCTGTGAAAGGGAAATCCCATGCTATGATTTGACATCAGGGTCCCGTTCTACGTG 1080
DB 1645 GGGCTGTGAAAGGGAAATCCCATGCTATGATTTGACATCAGGGTCCCGTTCTACGTG 1704
QY 1081 AGGGGCCCCAAACGTGAAAGCGGCTGTCTGAATCCCAACATGCTCTCAACATGACCTG 1140
DB 1705 AGGGGCCCCAAACGTGAAAGCGGCTGTCTGAATCCCAACATGCTCTCAACATGACCTG 1764
QY 1141 GCCCCACACCATCTGACACATTCAGGCGCTGGACATACCTTCGGATATGACGGGAATCC 1200
DB 1765 GCCCCACACCATCTGACACATTCAGGCGCTGGACATACCTTCGGATATGACGGGAATCC 1824
QY 1201 ATCTCTCAAGCTGTGACACAGGAGCGCGGTGAATCGGTTTCACTTGAAGAAAGATG 1260
DB 1825 ATCTCTCAAGCTGTGACACAGGAGCGCGGTGAATCGGTTTCACTTGAAGAAAGATG 1884
QY 1261 AGGCTCTGGCGGACCTCTTCTGTTGGAGAGGCGAAGCTGCTACACAGAGAGACAAT 1320
DB 1885 AGGCTCTGGCGGACCTCTTCTGTTGGAGAGGCGAAGCTGCTACACAGAGAGACAAT 1944
QY 1321 GACAAGGTGACGCCCCAGGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGACCTG 1380
DB 1945 GACAAGGTGACGCCCCAGGAGAGAACTTTCTGCCCAAGTACAGCGTGTGAAGACCTG 2004
QY 1381 TGTACGGTCTGAGTACACAGCGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440
DB 2005 TGTACGGTCTGAGTACACAGCGCGTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 2064
QY 1441 GAGGAGCCACAGGGGAAGCTGAAGCTGCAAGTCAAGGCGCCCATGCGGCTGGGCGCG 1500
DB 2065 GAGGAGCCACAGGGGAAGCTGAAGCTGCAAGTCAAGGCGCCCATGCGGCTGGGCGCG 2124
QY 1501 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTACGGGACAGGCGAGCGCTGCACC 1560
DB 2125 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTACGGGACAGGCGAGCGCTGCACC 2184
QY 1561 TGTGACAGCGGGGACTACAGCTCAGCTGCGCGGACCGCGGAAGAACTCTTCAAGAG 1620
DB 2185 TGTGACAGCGGGGACTACAGCTCAGCTGCGCGGACCGCGGAAGAACTCTTCAAGAG 2244
QY 1621 AAGTACAGGCCACAGCTATGTCCGACGTCCATCCGCTCAGTGGCCATCGAGGTGGAC 1680
DB 2245 AAGTACAGGCCACAGCTATGTCCGACGTCCATCCGCTCAGTGGCCATCGAGGTGGAC 2304
QY 1681 GGCAGGGTGTACCACTGAGCTGGTGTATGCGCCCGACCGGAAACCTCCACAGCGG 1740
DB 2305 GGCAGGGTGTACCACTGAGCTGGTGTATGCGCCCGACCGGAAACCTCCACAGCGG 2364
QY 1741 CATTCGCCAGGGGCCCTCTGAGGACCAAGATGACAGGATGGTGGGACCTTCAGTGGCACT 1800
DB 2365 CATTCGCCAGGGGCCCTCTGAGGACCAAGATGACAGGATGGTGGGACCTTCAGTGGCACT 2424
QY 1801 GAGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGACACATCGGTGCTACATC 1860
DB 2425 GAGAGGCTTCCCGACTACTCAGCGCGCAACCCCATTTAAAGTGACACATCGGTGCTACATC 2484
QY 1861 CTAGAGAACACACAGCTCAGTGTGACCTGTACAAGTCCCTGCAGGCGCTCGAAA 1920
DB 2485 CTAGAGAACACACAGCTCAGTGTGACCTGTACAAGTCCCTGCAGGCGCTCGAAA 2544
QY 1921 GACCACAGCTGCACATCGACCAAGATGTGAACCCCTGCAGAACAAATTTAAGAACCTG 1980
DB 2545 GACCACAGCTGCACATCGACCAAGATGTGAACCCCTGCAGAACAAATTTAAGAACCTG 2604

```
QY 1981 AGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCAGAGAAATGTGACTGTGCACAAAATC 2040
Db 2605 AGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCAGAGAAATGTGACTGTGCACAAAATC 2664
QY 2041 AGCTACACACCCAGCACAAGGCGGCTCAAGCACAGAGCTCCAGTCTGCATCCTTTC 2100
Db 2665 AGCTACACACCCAGCACAAGGCGGCTCAAGCACAGAGCTCCAGTCTGCATCCTTTC 2724
QY 2101 AGGAAGGCGCTGCAAGAGAAGCAAGGTGTGGTGTGTGGGAGCAGAGCGCAAGAAG 2160
Db 2725 AGGAAGGCGCTGCAAGAGAAGCAAGGTGTGGTGTGTGGGAGCAGAGCGCAAGAAG 2784
QY 2161 AAATCCCAAGCTGCTCAAGCGGCTGCGAACAAGACAGAGTCGACGATGCCAGGCTC 2220
Db 2785 AAATCCCAAGCTGCTCAAGCGGCTGCGAACAAGACAGAGTCGACGATGCCAGGCTC 2844
QY 2221 ACGTGCTTCAACCAGCAACACGACCTGGCAGACGGCGCTTCTGACACACTGGGSCCT 2280
Db 2845 ACGTGCTTCAACCAGCAACACGACCTGGCAGACGGCGCTTCTGACACACTGGGSCCT 2904
QY 2281 TTCTGTGCTGCACAGCGCCAAACAATAACAGTACTGTGTCATGAGGACCATCAATGAG 2340
Db 2905 TTCTGTGCTGCACAGCGCCAAACAATAACAGTACTGTGTCATGAGGACCATCAATGAG 2964
QY 2341 ACTCACAATTTCTCTCTGTAATTTGCACTTGGCTTCTAGAGTACTTTGATCTCAAC 2400
Db 2965 ACTCACAATTTCTCTCTGTAATTTGCACTTGGCTTCTAGAGTACTTTGATCTCAAC 3024
QY 2401 ACAGACCCCTACCGCTGATGAATGCGAGTGAACAACACTGGACAGGAGTGTCTTCAACAG 2460
Db 3025 ACAGACCCCTACCGCTGATGAATGCGAGTGAACAACACTGGACAGGAGTGTCTTCAACAG 3084
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGTTACAAGCAGTGTAAACCCCGG 2520
Db 3085 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGTTACAAGCAGTGTAAACCCCGG 3144
QY 2521 ACTCGAAACATGGACCTGGGACTTAAAGATGAGGAACTATGAGCAATACAGCAGTGT 2580
Db 3145 ACTCGAAACATGGACCTGGGACTTAAAGATGAGGAACTATGAGCAATACAGCAGTGT 3204
QY 2581 CAGCGTGAAGGTGCGCAGAAATGAAGAGACTTCTTCCAAATCACTGGACAACTGTGG 2640
Db 3205 CAGCGTGAAGGTGCGCAGAAATGAAGAGACTTCTTCCAAATCACTGGACAACTGTGG 3264
QY 2641 GAAGGCTGGAGGTTAAGAACACACAGAGTGGACCTCCAAACACATAGAGCATCAAC 2700
Db 3265 GAAGGCTGGAGGTTAAGAACACACAGAGTGGACCTCCAAACACATAGAGCATCAAC 3324
QY 2701 TGACTGCACAGGCAATGAAACCAATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3325 TGACTGCACAGGCAATGAAACCAATGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3384
QY 2761 AGGAGGCTGAGAAAGCAGACGACCTCTCAGTCAACATGACAGATTCGTGGAGGATAAC 2820
Db 3385 AGGAGGCTGAGAAAGCAGACGACCTCTCAGTCAACATGACAGATTCGTGGAGGATAAC 3444
QY 2821 CAGCAGGACGACAGATACTTCAGAGAGTCCATTTTGGCCCTGCTTTTCTGCTTTGATTA 2880
Db 3445 CAGCAGGACGACAGATACTTCAGAGAGTCCATTTTGGCCCTGCTTTTCTGCTTTGATTA 3504
QY 2881 TACTCTACCGCTGCACAAATGATTTTTCGTAATCAAAAGTCACCACTAACCCCTCC 2940
Db 3505 TACTCTACCGCTGCACAAATGATTTTTCGTAATCAAAAGTCACCACTAACCCCTCC 3564
QY 2941 CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3565 CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3064
QY 3001 CCCAAGGCGAAAGTCAATTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTGTTCTAAA 3060
Db 3625 CCCAAGGCGAAAGTCAATTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTGTTCTAAA 3684
```

```
QY 3061 TCTCTTATTTCTTTGGTTTGTCAAAAGAAAGAACTAAGAGCAGGACAGAGCAAGCT 3120
Db 3685 TCTCTTATTTCTTTGGTTTGTCAAAAGAAAGAACTAAGAGCAGGACAGAGCAAGCT 3744
QY 3121 GGAGAGGCTGAAACACAGTGCAGAGACGTTTGAATAGTCTAGTAGCACAAAAGAGATGA 3180
Db 3745 GGAGAGGCTGAAACACAGTGCAGAGACGTTTGAATAGTCTAGTAGCACAAAAGAGATGA 3804
QY 3181 CATTTACCTTAGCACTATAAAACCTGTTGCTCTCAAGAAACTGCTTTCATTGTATATAT 3240
Db 3805 CATTTACCTTAGCACTATAAAACCTGTTGCTCTCAAGAAACTGCTTTCATTGTATATAT 3863
QY 3241 GTGACTATTTTACATGATATCAACATGCGGAACTTTTATAGGGAACTTAATAGAAATCCCAA 3300
Db 3864 GTGACTATTTTACATGATATCAACATGCGGAACTTTTATAGGGAACTTAATAGAAATCCCAA 3923
QY 3301 TTTTTCAGAGTGTGCTGTCAATTAACCTCTGTGCGCAGTGTAAAGAAAGAAAAAAA 3360
Db 3924 TTTTTCAGAGTGTGCTGTCAATTAACCTCTGTGCGCAGTGTAAAGAAAGAAAAAAA 3983
QY 3361 AA 3362
Db 3984 AA 3985

RESULT 4
AX573191
LOCUS AX573191 4296 bp DNA linear PAT 29-NOV-2002
DEFINITION Sequence 13 from Patent WO02059327.
ACCESSION AX573191
VERSION AX573191.1 GI:26005109
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Rosen,S.D., Hemmerich,S., Tomita,M. and Palmeri,D.
TITLE Sulfatases and methods of use thereof
JOURNAL Patent: WO 02059327-A 13 01-AUG-2002;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (US)
FEATURES
Location/Qualifiers
1..4286
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1188 a 1145 c 1108 g 845 t
ORIGIN
Query Match 99.5%; Score 3346.8; DB 6; Length 4286;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3359; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 GACTCCCGCATCCCAAAAGAGCACCAGATCAGCAAAAAAGAGATGGGCCCCCGGAGC 60
Db 632 GACTCCCGCATCCCAAAAGAGCACCAGATCAGCAAAAAAGAGATGGGCCCCCGGAGC 691
QY 61 CTCGCTGTGTGCTGCTCTCGCAACTGTGTCTCCCTGCTGGTGGAGCTCGGCCCTTC 120
Db 692 CTCGCTGTGTGCTGCTCTCGCAACTGTGTCTCCCTGCTGGTGGAGCTCGGCCCTTC 751
QY 121 CTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAGAGGGACCGCAGGAAATCCGCCCCCAAC 180
Db 752 CTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAGAGGGACCGCAGGAAATCCGCCCCCAAC 811
QY 181 ATCATCTGTGTGCTGACGAGCAGCAGATGTGGAGCTGGGTTCCATGACAGTGTGTAAC 240
Db 812 ATCATCTGTGTGCTGACGAGCAGCAGATGTGGAGCTGGGTTCCATGACAGTGTGTAAC 871
QY 241 AAGACCCCGCGCATCATGAGCAGCGCGGGCGGCACTTTCATCAACGCTTCTGTCACACA 300
Db 872 AAGACCCCGCGCATCATGAGCAGCGCGGGCGGCACTTTCATCAACGCTTCTGTCACACA 931
```



```
Db 3092 CTACAGCTACAGCTCATGGAGCTGAGAGCTGCAAGGCTTACAGCAGTGTAAACCCCGG 3151
QY 2521 ACTGGAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATAGCAATACAGCGAGTTT 2580
Db 3152 ACTGGAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATAGCAATACAGCGAGTTT 3211
QY 2581 CAGCGTGGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCAGTCGGGACAACTGTGG 2640
Db 3212 CAGCGTGGAAAGTGGCCAGAAATGAAGAGACTTCTTCCAAATCAGTCGGGACAACTGTGG 3271
QY 2641 GAAGCTGGGAAGGTTTGAAGAAACACAGAGTGGACCTCCAAATAATGAGGAGCATACC 2700
Db 3272 GAAGCTGGGAAGGTTTGAAGAAACACAGAGTGGACCTCCAAATAATGAGGAGCATACC 3331
QY 2701 TGACTGCACAGGCAATGAAAGCAATGAGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3332 TGACTGCACAGGCAATGAAAGCAATGAGTGGTGTATTTCCAGCAGACCTGTGCTATTGGCC 3391
QY 2761 AGGAGGCTGAGAAGCAAGCACCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC 2820
Db 3392 AGGAGGCTGAGAAGCAAGCACCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC 3451
QY 2821 CAGCAGGACAGATTAACCTTCAGAGAGTCCATTTTGGCCCTGCTTTTCTGCTTGGATTA 2880
Db 3452 CAGCAGGACAGATTAACCTTCAGAGAGTCCATTTTGGCCCTGCTTTTCTGCTTGGATTA 3511
QY 2881 TACCTCACAGCTGCACAAATGCAATTTTTCGATCAAAAGTCAACCACTAACCCCTCCC 2940
Db 3512 TACCTCACAGCTGCACAAATGCAATTTTTCGATCAAAAGTCAACCACTAACCCCTCCC 3571
QY 2941 CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3572 CCAGAGCTCAAAAGGAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3060
QY 3001 CCCAAGGCGAAAGTCAATTGGAAATTTTAAATCATAGGGAAAGCAGTCTCTTCTTAAA 3060
Db 3632 CCCAAGGCGAAAGTCAATTGGAAATTTTAAATCATAGGGAAAGCAGTCTCTTCTTAAA 3091
QY 3061 TCCTCTTATCTTTTGGTTTCTCAAGAAAGAACTTAAGAGCAGACAGAGCGCAACGT 3120
Db 3692 TCCTCTTATCTTTTGGTTTCTCAAGAAAGAACTTAAGAGCAGACAGAGCGCAACGT 3180
QY 3121 GGAGAGGCTGAAACAGGTGCAGAGAGCTTTCACAAATGAGTCAGTACAGTACAGAAAGAGATGA 3180
Db 3752 GGAGAGGCTGAAACAGGTGCAGAGAGCTTTCACAAATGAGTCAGTACAGTACAGAAAGAGATGA 3811
QY 3181 CATTTACCTAGCATTAAACCTCTGTTTCTCTGAAAGAACTCCCTTCATTGTATATAT 3240
Db 3812 CATTTACCTAGCA-TATAAACCTCTGTTTCTCTGAAAGAACTCCCTTCATTGTATATAT 3870
QY 3241 GTGACTATTACATGTATACATGGGAACTTTTAGGGNAACCTATAGAAATCCCAA 3300
Db 3871 GTGACTATTACATGTATACATGGGAACTTTTAGGGNAACCTATAGAAATCCCAA 3390
QY 3301 TTTTCAGAGTGGTGGTCTCAATAAACCTCTGTGGCAGTGTAAAGAAAGAAAAA 3360
Db 3991 TTTTCAGAGTGGTGGTCTCAATAAACCTCTGTGGCAGTGTAAAGAAAGAAAAA 3990
QY 3361 AA 3362
Db 3991 AA 3992
```

```
RESULT 5
AX211371
LOCUS AX211371 3348 bp DNA linear PAT 06-SEP-2001
DEFINITION Sequence 17 from Patent WO0159127.
ACCESSION AX211371
VERSION AX211371.1 GI:15523740
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS Tang,Y.T., Yue,H., Baughn,M.R., Yao,M.G., Bandman,O., Azimzai,Y.,
Lai,P., Gandhi,A.R., Ring,H.Z., Shih,L.D., Yang,J. and Pollocky,J.L.
TITLE Drug metabolizing enzymes
JOURNAL Patent: WO 0159127-A 17 16-AUG-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..3348
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1558210CB1"
BASE COUNT 927 a 933 c 869 g 619 t
ORIGIN
Query Match 99.5%; Score 3344.8; DB 6; Length 3348;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 12 CCCAAAAGAACCCAGATCAGCAAAAAAGAGATGGGCCCCCGAGCCCTCGTGTGTG 71
Db 1 CCCAAAAGAACCCAGATCAGCAAAAAAGAGATGGGCCCCCGAGCCCTCGTGTGTG 60
QY 72 CTTGCTGTCGGCAACTGTGTTCTCCTGCTGGTGGAGACTCGGCTTCTTGTGACCA 131
Db 61 CTTGCTGTCGGCAACTGTGTTCTCCTGCTGGTGGAGACTCGGCTTCTTGTGACCA 120
QY 132 CGGCTGAAAGGACAGTTTCAGAGGAGCGCAGGAAACATCGCCCAACATCATCTGCT 191
Db 121 CGGCTGAAAGGACAGTTTCAGAGGAGCGCAGGAAACATCGCCCAACATCATCTGCT 180
QY 192 GTGACGAGCAGACAGAGATGTGGAGTCTGGTTCATGAGTGTATGAACAAGCCGGCG 251
Db 181 GCTGACGAGCAGACAGAGATGTGGAGTCTGGTTCATGAGTGTATGAACAAGCCGGCG 240
QY 252 CATCATGAGCAGAGCGCGCGCGCACTTCATCAAGCCCTTCGTGACCAACCATGTGTG 311
Db 241 CATTTAGAGCAGGCGGGGCGCGCACTTCATCAAGCCCTTCGTGACCAACCATGTGTG 300
QY 312 CCCCTCAGCTCTCCATCTCTCAGTGGCAAGTAGTCCAAACCAACAACTACACAA 371
Db 301 CCCCTCAGCTCTCCATCTCTCAGTGGCAAGTAGTCCAAACCAACAACTACACAA 360
QY 372 CAATGAGAACTGCTCTCGCCCTCTGGCAGGCAAGCAGAGCGCGCACTTTGCCGT 431
Db 361 CAATGAGAACTGCTCTCGCCCTCTGGCAGGCAAGCAGAGCGCGCACTTTGCCGT 420
QY 432 GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGATACAA 491
Db 421 GTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGATACAA 480
QY 492 CGGCTCTCAAGTGGCCACCGCGCTGGAAGAGTGGTGGAGCTCTTAAATACTCCGCTT 551
Db 481 CGGCTCTCAAGTGGCCACCGCGCTGGAAGAGTGGTGGAGCTCTTAAATACTCCGCTT 540
QY 552 TTATAACTACAGCTGTGTGGAAAGGCTGAAAGAAAGCAAGCAGCTCCGACTACTCAA 611
Db 541 TTATAACTACAGCTGTGTGGAAAGGCTGAAAGAAAGCAAGCAGCTCCGACTACTCAA 600
QY 612 GGATTACCTTCACAGACCTCATCAATGACAGCGTGTAGCTTCTTCCGCAAGTCCAGAA 671
Db 601 GGATTACCTTCACAGACCTCATCAATGACAGCGTGTAGCTTCTTCCGCAAGTCCAGAA 660
QY 672 GATGTACCGGACAGCGCAGCTCTCATGTGTATCAGCCATCAGCCCCCGCGCCCTGA 731
Db 661 GATGTACCGGACAGCGCAGCTCTCATGTGTATCAGCCATCAGCCCCCGCGCCCTGA 720
QY 732 GGATTACGCCCCCAATATTCAGGCTCTTCCAAACGCACTCAGCAGCATCACGCCGAG 791
Db 721 GGATTACGCCCCCAATATTCAGGCTCTTCCAAACGCACTCAGCAGCATCACGCCGAG 780
QY 792 CTACAACTACGCGCCCAACCCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGNA 851
```

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1
Tang,Y.T., Yue,H., Baughn,M.R., Yao,M.G., Bandman,O., Azimzai,Y.,
Lai,P., Gandhi,A.R., Ring,H.Z., Shih,L.D., Yang,J. and Pollocky,J.L.
Drug metabolizing enzymes
Patent: WO 0159127-A 17 16-AUG-2001;
Incyte Genomics, Inc. (US)

Location/Qualifiers
1..3348
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1558210CB1"

BASE COUNT 927 a 933 c 869 g 619 t

ORIGIN

Db 781 CTACAACTACGGGCCCAACCCGGAACAAACCTGGATCATGGCTACACGGGGCCCATGAA 840
QY 852 GCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGACCTCATGTC 911
Db 841 GCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGACCTCATGTC 900
QY 912 GGTGGACGACTCCATGGAGACGATTTACAAATGCTGTTGAGACGGCGGAGCTGGACAA 971
Db 901 GGTGGACGACTCCATGGAGACGATTTACAAATGCTGTTGAGACGGCGGAGCTGGACAA 960
QY 972 CAGGTACATCGGTATACACCGCGACACAGGTTACACATCGGCCAGTTTGGCCGTGGAA 1031
Db 961 CAGGTACATCGGTATACACCGCGACACAGGTTACACATCGGCCAGTTTGGCCGTGGAA 1020
QY 1032 AGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCGGTTCTACGTGAGGGGCCCAA 1091
Db 1021 AGGGAATCCATGCCATATGAGTTTGACATCAGGGTCCGGTTCTACGTGAGGGGCCCAA 1080
QY 1092 CGTGGAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCCCCACCAT 1151
Db 1081 CGTGGAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTTGACCTGGCCCCCACCAT 1140
QY 1152 CTTGGACATTTGAGGCTGGACATACCTCGGATATGAGCGGAAATCCATCTCAAGCT 1211
Db 1141 CTTGGACATTTGAGGCTGGACATACCTCGGATATGAGCGGAAATCCAATCTCAAGCT 1200
QY 1212 GCTGGACACGGAGCGGCTGGATCGGTTTCACTTTGAAAAAGAGATGAGGGTCTGGCG 1271
Db 1201 GCTGGACACGGAGCGGCTGGATCGGTTTCACTTTGAAAAAGAGATGAGGGTCTGGCG 1260
QY 1272 GGAATCTCTTCTGGTGGAGAGGCAAGCTGTACACAGAGAGACAATGACAGAGTGA 1331
Db 1261 GGAATCTCTTCTGGTGGAGAGGCAAGCTGTACACAGAGAGACAATGACAGAGTGA 1320
QY 1332 CGCCAGGAGAGAAATTTCTGCCCCAAGTACAGCGTGTGAAGGACCTGTGTCAGCGTGC 1391
Db 1321 CGCCAGGAGAGAAATTTCTGCCCCAAGTACAGCGTGTGAAGGACCTGTGTCAGCGTGC 1380
QY 1392 TGAGTACCAAGCGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGGAGGAGCGCCAC 1451
Db 1381 TGAGTACCAAGCGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGGAGGAGCGCCAC 1440
QY 1452 GGGGAAGCTGAAGCTGCATAAGTGAAGGCGCCCATGCGGCTGGGCGGCGAGCAGAGCCCT 1511
Db 1441 GGGGAAGCTGAAGCTGCATAAGTGAAGGCGCCCATGCGGCTGGGCGGCGAGCAGAGCCCT 1500
QY 1512 CTCGAACCTCGTGGCCAGTACTAGCGGCGAGGCGGAGGCGCTGCACTGTGACAGCGG 1571
Db 1501 CTCGAACCTCGTGGCCAGTACTAGCGGCGAGGCGGAGGCGCTGCACTGTGACAGCGG 1560
QY 1572 GGAATCAAGCTCAGCTGGCGGACCGCGGAAAAAATCTTCAAGAGAAAGTACAAGGC 1631
Db 1561 GGAATCAAGCTCAGCTGGCGGACCGCGGAAAAAATCTTCAAGAGAAAGTACAAGGC 1620
QY 1632 CAGTATGTCGAGTGGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGGAGGTGTA 1691
Db 1621 CAGTATGTCGAGTGGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGGAGGTGTA 1680
QY 1692 CCAGTGGGCTGGGTGATCGGCCCGAGCCCGAACCCTCACCAAGCGCATCTGGCCAGG 1751
Db 1681 CCAGTGGGCTGGGTGATCGGCCCGAGCCCGAACCCTCACCAAGCGCATCTGGCCAGG 1740
QY 1752 GGCCCTTGAGGACCAAGATGACAGGATGGTGGGACCTTCAGTGGCACTTGGAGGCTTCC 1811
Db 1741 GGCCCTTGAGGACCAAGATGACAGGATGGTGGGACCTTCAGTGGCACTTGGAGGCTTCC 1800
QY 1812 CGACTACTAGCGCCCAACCCCATTAAGTGAACATCGGTGTACATCTTAGAGAGCA 1871
Db 1801 CGACTACTAGCGCCCAACCCCATTAAGTGAACATCGGTGTACATCTTAGAGAGCA 1860
QY 1872 CACAGTCCAGTGTGACCTGGACCTGTACAGTCCCTGCAAGCTGGAGGCTGGAAAGACCAAGCT 1931

Db 1861 CACAGTCCAGTGTGACCTGGACCTGTCAAGTCCCTGCAAGGCTGGAAAGACCAAGCT 1920
QY 1932 GCACATCCAGCCACGAGATTTGAAACCTCTGCAGAACAAAAATTAAGAACTTGGAGGAGTCCG 1991
Db 1921 GCACATCCAGCCACGAGATTTGAAACCTCTGCAGAACAAAAATTAAGAACTTGGAGGAGTCCG 1980
QY 1992 AGTGTACCTGTGAAGAAAAAGCGGCGAGAGAAATGTGACTGTCAAAAATCAGCTACACAC 2051
Db 1981 AGTGTACCTGTGAAGAAAAAGCGGCGAGAGAAATGTGACTGTCAAAAATCAGCTACACAC 2040
QY 2052 CCAGCAAAAGGCGGCTCAAGCAGACAGGCTCCAGTCTGCAATCTTTTCAGGAAGGCGCT 2111
Db 2041 CCAGCAAAAGGCGGCTCAAGCAGACAGGCTCCAGTCTGCAATCTTTTCAGGAAGGCGCT 2100
QY 2112 GCAGAGAAAGCAAGGTGTGGCTGTGCGGAGAGAGAGCGCAAGAGAAATCTCGGCAA 2171
Db 2101 GCAGAGAAAGCAAGGTGTGGCTGTGCGGAGAGAGAGCGCAAGAGAAATCTCGGCAA 2160
QY 2172 GCTGCTCAAGCGGCTCAGAAACCAACACACAGCTGCAGCTGCCAGGCTCACCGTCTTAC 2231
Db 2161 GCTGCTCAAGCGGCTCAGAAACCAACACACAGCTGCAGCTGCCAGGCTCACCGTCTTAC 2220
QY 2232 CCAGCAAAACAGCACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGTGCGCTG 2291
Db 2221 CCAGCAAAACAGCACTGGCAGACGGCGCTTTCTGGACACTGGGCGCTTTCTGTGCGCTG 2280
QY 2292 CACGAGCGCCAAACAAATACACGTACTGTGTCATGAGGACCATCAATGAGACTCACAAAT 2351
Db 2281 CACGAGCGCCAAACAAATACACGTACTGTGTCATGAGGACCATCAATGAGACTCACAAAT 2340
QY 2352 CTTCTTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTA 2411
Db 2341 CTTCTTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAACACAGACCCCTA 2400
QY 2412 CCACTGATGATGCAAGTGAACACACTGGACAGGATGTCTCAACACAGTACACGTACA 2471
Db 2401 CCACTGATGATGCAAGTGAACACACTGGACAGGATGTCTCAACACAGTACACGTACA 2460
QY 2472 GCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGGACTCGAAACAT 2531
Db 2461 GCTCATGGAGCTGAGGAGCTGCAAGGTTTACAGCAGTGTAAACCCCGGACTCGAAACAT 2520
QY 2532 GGAACCTGGGACTTAAAGATGGAGAAAGCTATAGCAATACAGGAGCTTTTCAGGCTGAAA 2591
Db 2521 GGAACCTGGGACTTAAAGATGGAGAAAGCTATAGCAATACAGGAGCTTTTCAGGCTGAAA 2580
QY 2592 GTGGCCAGAAATGAGAGACCTTTCTCAATCACTGGGCAAACTGTGGGAAGGCTGGGA 2651
Db 2581 GTGGCCAGAAATGAGAGACCTTTCTCAATCACTGGGCAAACTGTGGGAAGGCTGGGA 2640
QY 2652 AGGTTAAGAAACCAACAGAGGTGGACCTCCAAAAACATAGAGGATCACTGTGACACAG 2711
Db 2641 AGGTTAAGAAACCAACAGAGGTGGACCTCCAAAAACATAGAGGATCACTGTGACACAG 2700
QY 2712 GCATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGA 2771
Db 2701 GCATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCTGA 2760
QY 2772 GAAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAACCCAGCAGAGCA 2831
Db 2761 GAAAGCAAGCAACGCACTCTCAGTCAACATGACAGATTTCTGGAGGATAACCCAGCAGAGCA 2820
QY 2832 GAGATACTTCAGGAAGTCCATTTTTCGCTATCAAAAAGTCAACCTTAAGCCCTCCCGAGAGCTCA 2891
Db 2821 GAGATACTTCAGGAAGTCCATTTTTCGCTATCAAAAAGTCAACCTTAAGCCCTCCCGAGAGCTCA 2880
QY 2892 CTGCACAAAAATGCATTTTTCGCTATCAAAAAGTCAACCTTAAGCCCTCCCGAGAGCTCA 2951
Db 2881 CTGCACAAAAATGCATTTTTCGCTATCAAAAAGTCAACCTTAAGCCCTCCCGAGAGCTCA 2940
QY 2952 CAAAGGAAAAACGAGAGAGCGAGAGAGATTTTCTTGGAAATTTCTCCCAAGGGCGGA 3011
Db 2941 CAAAGGAAAAACGAGAGAGCGAGAGAGATTTTCTTGGAAATTTCTCCCAAGGGCGGA 3000

QY 3012 AAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAATCCTCTATTTC 3071
Db
QY 3001 AAGTCATTGGAAATTTTAAATCATAGGGGAAAGCAGTCTGTTCTAAATCCTCTATTTC 3060
Db
QY 3072 TTTTGGTTGTGCACAAAGAAAGAACTAAGAACAGGAGGACAGAGGAGGCTGA 3131
Db
QY 3061 TTTTGGTTGTGCACAAAGAAAGAACTAAGAACAGGAGGACAGAGGAGGCTGA 3120
QY 3132 AAACAGTCAGAGAGCTTTTGACAAATGAGTCAGTAGCACAAGAGAGATGACATTTTACCTAG 3191
Db
QY 3121 AAACAGTCAGAGAGCTTTTGACAAATGAGTCAGTAGCACAAGAGAGATGACATTTTACCTAG 3180
QY 3192 CACTATAAACCTGTGTGCTCTGAAAGAACTGCTTCATTTGATATATATGACTATTTA 3251
Db
QY 3181 CACTATAAACCTGTGTGCTCTGAAAGAACTGCTTCATTTGATATATATGACTATTTA 3240
QY 3252 CATGTAATCAACATGGGAACCTTTTAGGGGAACTTAATAAGAAATCCCAATTTTTCAGGAGT 3311
Db
QY 3241 CATGTAATCAACATGGGAACCTTTTAGGGGAACTTAATAAGAAATCCCAATTTTTCAGGAGT 3300
QY 3312 GGTGTGTCAATAAACGCTGTGTGCGCCAGTGTAAAGAAATTTTAAAAA 3359
Db
QY 3301 GGTGTGTCAATAAACGCTGTGTGCGCCAGTGTAAAGAAATTTTAAAAA 3348

RESULT 6
LOCUS AR252433 3906 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 83 from patent US 6478825.
ACCESSION AR252433
VERSION AR252433.1 GI:27300341
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3906)
AUTHORS Winterbottom J.M., Boyce, T.M. and Xaes, D.
TITLE Implant, method of making same and use of the implant for the treatment of bone defects
JOURNAL Patent: US 6478825-A 83 12-NOV-2002;
FEATURES
Location/Qualifiers
source 1..3906
/organism="unknown"
BASE COUNT 1047 a 1089 c 1041 g 729 t
ORIGIN

Query Match 98.8%; Score 3323.4; DB 6; Length 3906;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 GACTCCCGCATCCCAAAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 60
Db
QY 563 GACTCCCGCATCCCAAAAGAGACACAGATCAGCAAAAAGAGATGGGCCCCCGAGC 622
QY 61 CTCGTGTGTGTGCTGTGTCGCAACTGTGTCTCCCTGTGGTGGAGCTCGGCTTC 120
Db
QY 623 CTCGTGTGTGTGCTGTGTCGCAACTGTGTCTCCCTGTGGTGGAGCTCGGCTTC 682
QY 121 CTGTGCGACACCCGCTTAAGAGGAGGTTTTCAGAGGGACCGCAGACATCCGCCCAAC 180
Db
QY 693 CTGTGCGACACCCGCTTAAGAGGAGGTTTTCAGAGGGACCGCAGACATCCGCCCAAC 742
QY 181 ATCATCTCTGTGCTGACGAGACGACAGGATGTGGAGTGGGTTCCATGCAAGTGATGAAC 240
Db
QY 743 ATCATCTCTGTGCTGACGAGACGACAGGATGTGGAGTGGGTTCCATGCAAGTGATGAAC 802
QY 241 AAGACCGGCGCATCATGAGACGAGGCGGGCGGCACATTCATCAACGCTTCGTGACACA 300
Db
QY 803 AAGACCGGCGCATCATGAGACGAGGCGGGCGGCACATTCATCAACGCTTCGTGACACA 862
QY 301 CCCATGTCTGCTGCTCAGCTCTCCATCTCTCACTGTGCAAGTACGTCCCAACCAAC 360

Db 863 CCATGTGTGCTGCTCAGCTCTCTCCATCTCACTGGCAAGTACGTCCCAACCAACAAAC 922
QY 361 ACTTACACCAACAAATGAGAACTGCTCTCGCTCTGTCAGGACACAGACAGAGCCGC 420
Db 923 ACTTACACCAACAAATGAGAACTGCTCTCGCTCTGTCAGGACACAGACAGAGCCGC 982
QY 421 ACTTTTGGCTGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 480
Db 983 ACTTTTGGCTGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTT 1042
QY 481 AATGATACAAAGGCTCTTACCTGCGACCCGCTGGAAGAGTGGTGGACTCTCTTAA 540
Db 1043 AATGATACAAAGGCTCTTACCTGCGACCCGCTGGAAGAGTGGTGGACTCTCTTAA 1102
QY 541 AACTCCGCTTTTATAAATTACACAGCTGTGTGGAACGGGTTGAAAGAAAGACCGCTCC 600
Db 1103 AACTCCGCTTTTATAAATTACACAGCTGTGTGGAACGGGTTGAAAGAAAGACCGCTCC 1162
QY 601 GACTTCTCAAGGATTAACCTCAAGACCTCATCAATGACAGCTGAGCTTCTTCGCG 660
Db 1163 GACTTCTCAAGGATTAACCTCAAGACCTCATCAATGACAGCTGAGCTTCTTCGCG 1222
QY 661 AGTCTCAAGAGATGTACCCGACAGGCGAGTCTCATGCTCATGAGCATCAGCCGCC 720
Db 1223 AGTCTCAAGAGATGTACCCGACAGGCGAGTCTCATGCTCATGAGCATCAGCCGCC 1282
QY 721 CACGCGCTGAGGATTAACGCCCCACAAATATTACAGCTCTTTCCAAACGCAFTCTCAGCAC 780
Db 1283 CACGCGCTGAGGATTAACGCCCCACAAATATTACAGCTCTTTCCAAACGCAFTCTCAGCAC 1342
QY 781 ATCAGCGCGAGTCAAACTAGCGCGCCCAACCGGACAAACACTGATCATGCTTACAGG 840
Db 1343 ATCAGCGCGAGTCAAACTAGCGCGCCCAACCGGACAAACACTGATCATGCTTACAGG 1402
QY 841 GGGCCCATGAAGCCCATCCACATGGAATTTCCAAACATGCTCCAGCGGAAGCGTTGCGAG 900
Db 1403 GGGCCCATGAAGCCCATCCACATGGAATTTCCAAACATGCTCCAGCGGAAGCGTTGCGAG 1462
QY 901 ACCCTCATGTGCGTGGAGCTTCCATGAGAGAGATTTACAACTGCTGTTGAGACGGGC 960
Db 1463 ACCCTCATGTGCGTGGAGCTTCCATGAGAGAGATTTACAACTGCTGTTGAGACGGGC 1522
QY 961 GAGCTGGGACAAACGCTATACACCGCGGACCGTTCACCATGCTCCACATCGGCAGTTT 1020
Db 1523 GAGCTGGGACAAACGCTATACACCGCGGACCGTTCACCATGCTCCACATCGGCAGTTT 1582
QY 1021 GGCCTGTGTAAGGAAATTCATGCTATGAGTTTGAATCAGGCTCCGCTTCTTACGTTG 1080
Db 1583 GGCCTGTGTAAGGAAATTCATGCTATGAGTTTGAATCAGGCTCCGCTTCTTACGTTG 1642
QY 1081 AGGGCCCCAACGCTGGAAGCGGCTGCTGAATCCCACTGCTCCCAATTCAGACCTG 1140
Db 1643 AGGGCCCCAACGCTGGAAGCGGCTGCTGAATCCCACTGCTCCCAATTCAGACCTG 1702
QY 1141 GCCCCCAACATCTCGACATTCAGGCTGAGCATACCTGCGGATATGAGCGGAAATTC 1200
Db 1703 GCCCCCAACATCTCGACATTCAGGCTGAGCATACCTGCGGATATGAGCGGAAATTC 1762
QY 1201 ATCTCTCAAGCTGCTGACACAGGCGGCGGCTGATCGGTTTCACTTGAAGAGAGATG 1260
Db 1763 ATCTCTCAAGCTGCTGACACAGGCGGCGGCTGATCGGTTTCACTTGAAGAGAGATG 1822
QY 1261 AGGGCTCTGCGGGAATCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACACAGAGAGCAAT 1320
Db 1823 AGGGCTCTGCGGGAATCTCTTCTTGGTGAGAGAGGCAAGCTGTCTACACAGAGAGCAAT 1882
QY 1321 GACAGGTGGAGCGCCCGAGAGAGAGAACTTTCTGCCAAGTACACGCTGTGAGAGACCTG 1380
Db 1883 GACAGGTGGAGCGCCCGAGAGAGAGAACTTTCTGCCAAGTACACGCTGTGAGAGACCTG 1942
QY 1381 TGTGAGCGTGTGCTGAGTACAGACGCGGCTGTGAGCAGCTGGGACAGAGAGTGGCAGTGTG 1440
Db 1943 TGTGAGCGTGTGCTGAGTACAGACGCGGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTG 2002

1441 GAGGAGCCACCGGGAAGCTGAAGCTGCATAAGTGCAGAGGCGCCCATCGGCTGGCGGC 1500
Db
2003 GAGGAGCCACCGGGAAGCTGAAGCTGCATAAGTGCAGAGGCGCCCATCGGCTGGCGGC 2062
Qy 1501 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGGCGCAGGCGAGGCGCTGCACC 1560
Db 2063 AGCAGAGCCCTCTCCAACTCTGTCGCCAAGTACTAGGCGCAGGCGAGGCGCTGCACC 2122
Qy 1561 TGTGACAGCGGGGACTTACAACTCTGAGCTGGCGGACGCGCGGAAAAAACTCTTCAAGAAG 1620
Db 2123 TGTGACAGCGGGGACTTACAACTCTGAGCTGGCGGACGCGCGGAAAAAACTCTTCAAGAAG 2182
Qy 1621 AAGTACAGGCGCAGCTATGTCGCGAGTGCCTCCATCGCTCCAGTGGCCCATGAGTGGAC 1680
Db 2183 AAGTACAGGCGCAGCTATGTCGCGAGTGCCTCCATCGCTCCAGTGGCCCATGAGTGGAC 2242
Qy 1681 GGCAGGGTGTACACGTAGGCTGGGTGATGCGGCCAGCGCCGAAAACTCACCAGCGG 1740
Db 2243 GGCAGGGTGTACACGTAGGCTGGGTGATGCGGCCAGCGCCGAAAACTCACCAGCGG 2302
Qy 1741 CACTGCGCAGGCGCCCTGAGACCAAGATGACAAAGATGTTGGGACTTCAGTGGCACT 1800
Db 2303 CACTGCGCAGGCGCCCTGAGACCAAGATGACAAAGATGTTGGGACTTCAGTGGCACT 2362
Qy 1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAAGTGACACATCGGTGTACATC 1860
Db 2363 GGAGGCTTCCCGACTACTCAGCGCCCAACCCATTAAAGTGACACATCGGTGTACATC 2422
Qy 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGACGCGCTGGAAA 1920
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTATCAAGTCCCTGACGCGCTGGAAA 2482
Qy 1921 GACCACAGCTGCATGACACAGAGATGGAACCTTCGACAGACAAATTAAGAACCTG 1980
Db 2483 GACCACAGCTGCATGACACAGAGATGGAACCTTCGACAGACAAATTAAGAACCTG 2542
Qy 1981 AGGGAAGTCCGAGGTCACTGGAAGAAAAAGCGGCGCAGAGAAATGTGCTGCACAAAATC 2040
Db 2543 AGGGAAGTCCGAGGTCACTGGAAGAAAAAGCGGCGCAGAGAAATGTGCTGCACAAAATC 2602
Qy 2041 AGGTACACACCGACGACAAAGCGGCTTCAGGACAGAGGCTTCAGTCTGATCTCTTTC 2100
Db 2603 AGGTACACACCGACGACAAAGCGGCTTCAGGACAGAGGCTTCAGTCTGATCTCTTTC 2662
Qy 2101 AGGAAGGCGCTGCAAGAGAAAGACAGGTGTGGCTGTTGCGGAGCAGAGCGCAAGAG 2160
Db 2663 AGGAAGGCGCTGCAAGAGAAAGACAGGTGTGGCTGTTGCGGAGCAGAGCGCAAGAG 2722
Qy 2161 AAATCCGCAAGCTGCTCAAGCGCTTCAGAGACAAAGACAGTGCAGATGCCAGGCTC 2220
Db 2723 AAATCCGCAAGCTGCTCAAGCGCTTCAGAGACAAAGACAGTGCAGATGCCAGGCTC 2782
Qy 2221 AGTGTCTTCCACCGACACAGCAGCTGCGAGAGCGGCGCTTCTGACACATGGGCGCT 2280
Db 2783 AGTGTCTTCCACCGACACAGCAGCTGCGAGAGCGGCGCTTCTGACACATGGGCGCT 2842
Qy 2281 TTCTGTGCTGACCGCCCAACCAATTAACGTACTGTGTGATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGACCGCCCAACCAATTAACGTACTGTGTGATGAGGACCATCAATGAG 2902
Qy 2341 ACTCACAATTTCTCTCTGTGATTTGGAATGCGCTTCCTAGTACTTTGATCTCAAC 2400
Db 2903 ACTCACAATTTCTCTCTGTGATTTGGAATGCGCTTCCTAGTACTTTGATCTCAAC 2962
Qy 2401 ACAGACCCCTACAGCTGATGATGAGTGAAGTGAACACACTGACAGGAGTGTCTCAACAG 2460
Db 2963 ACAGACCCCTACAGCTGATGATGAGTGAAGTGAACACACTGACAGGAGTGTCTCAACAG 3022
Qy 2461 CTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTTCAAGAGTGTAAACCCCGG 2520
Db 3023 CTACAGTACAGCTCATGAGCTGAGGAGCTGCAAGGGTTTCAAGAGTGTAAACCCCGG 3082

Qy 2521 ACTCGAAACATGACCTGGGACTTAAAGATGAGGAGAGCTATGAGCAATACAGGCGATTT 2580
Db 3083 ACTCGAAACATGACCTG-----GATGAGGAGAGCTATGAGCAATACAGGCGATTT 3133
Qy 2581 CAGGCTCGAAGTGGCCAGAAATGAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 2640
Db 3134 CAGGCTCGAAGTGGCCAGAAATGAGAGACCTTCTTCCAAATCACTGGGACAACTGTGG 3193
Qy 2641 GAAGGCTGGGAAGGTTTAAAGAAACACAGAGGTGGACCTCCAAAAACATAGAGCATCAC 2700
Db 3194 GAAGGCTGGGAAGGTTTAAAGAAACACAGAGGTGGACCTCCAAAAACATAGAGCATCAC 3253
Qy 2701 TGACTGCAAGGCAATGAAAAACCATGTGGTGAATTCAGCAGACCTGTGTATTGGCC 2760
Db 3254 TGACTGCAAGGCAATGAAAAACCATGTGGTGAATTCAGCAGACCTGTGTATTGGCC 3313
Qy 2761 AGAGGCTGAGAAACCAAGCAGCAGCTCTCAGTCAACATGACAGATTCGAGAGATAAC 2820
Db 3314 AGAGGCTGAGAAACCAAGCAGCAGCTCTCAGTCAACATGACAGATTCGAGAGATAAC 3373
Qy 2821 CAGCAGAGCAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTGGATTA 2880
Db 3374 CAGCAGAGCAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGGCTTGGATTA 3433
Qy 2881 TACCTCAGCAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCTCCC 2940
Db 3434 TACCTCAGCAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCAACACTAACCTCCC 3493
Qy 2941 CCAGAAGCTCACAAGGAAAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAAGCTCACAAGGAAAAACGAGAGAGCGAGAGAGATTTCTTGGAAATTTCT 3553
Qy 3001 CCCAAGGCGGAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAAGCAGTCTGTCTAAA 3060
Db 3554 CCCAAGGCGGAAGTCAATTTGGAATTTTAAATCATAGGGGAAAAAGCAGTCTGTCTAAA 3613
Qy 3061 TCCTCTATTCTTTTGGTTTGTCAAAAAGAGAACTAAGAAGCAGGACAGAGCAACGT 3120
Db 3614 TCCTCTATTCTTTTGGTTTGTCAAAAAGAGAACTAAGAAGCAGGACAGAGCAACGT 3673
Qy 3121 GGAGAGCTGMAACAGTCAGAGAGGTTTGAATGATGATGATGATGATGATGATGATGATG 3180
Db 3674 GGAGAGCTGMAACAGTCAGAGAGGTTTGAATGATGATGATGATGATGATGATGATGATG 3733
Qy 3181 CATTTACCTAGCACTATAAACCTTGTCTGCTCTGAAGAACTGCTTCAATGTATATAT 3240
Db 3734 CATTTACCTAGCACTATAAACCTTGTCTGCTCTGAAGAACTGCTTCAATGTATATAT 3793
Qy 3241 GTGACTATTATCATGTAATCAATGGAACCTTTTGAAGGAACTTAATGAAGAAATCCCAA 3300
Db 3794 GTGACTATTATCATGTAATCAATGGAACCTTTTGAAGGAACTTAATGAAGAAATCCCAA 3853
Qy 3301 TTTTTCAGGAGTGGTGTGTCATTAACCGCTCTGTGCGCAGTGTAAAGAAAAA 3353
Db 3854 TTTTTCAGGAGTGGTGTGTCATTAACCGCTCTGTGCGCAGTGTAAAGAAAAA 3906

RESULT 7

AX403196 LOCUS 3906 bp DNA linear PAT 14-JUN-2002
Sequence 83 from Patent WO0073454.
AX403196 ACCESSION
AX403196.1 GI:21436787
VERSION
KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Ashkenazi, A.J., Baker, K.P., Botstein, D., Desnovers, L., Eaton, D.,
Ferrara, N., Gerber, H., Gerritsen, M., Goddard, A., Godowski, P.,
Grimaldi, C.J., Gurney, A.L., Kljavin, I., Napier, M.A., Pan, J.,
Paoni, N.F., Roy, M., Stewart, T.A., Tumas, D., Watanabe, C.K.,

Williams, P., Wood, W.I. and Zhang, Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0073454-A 83 07-DEC-2000;
Genentech Inc. (US)

FEATURES
source
Location/Qualifiers
1..3906
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1047 a 1089 c 1041 g 729 t
ORIGIN

Query Match 98.8%; Score 3323.4; DB 6; Length 3906;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
QY 1 GACTCCCGCATCCCAAGAGCAGACAGATCAGCAAAAAGAGATGGGCCCGGAGC 60
DB 563 GACTCCCGCATCCCAAGAGCAGACAGATCAGCAAAAAGAGATGGGCCCGGAGC 622
QY 61 CTGCTGTGTCTTGTCTGTCGCAACTGTGTCTTCTCTGCTGGGTGGAAGCTTCGGCTTC 120
DB 623 CTGCTGTGTCTTGTCTGTCGCAACTGTGTCTTCTCTGCTGGGTGGAAGCTTCGGCTTC 682
QY 121 CTGTGCGACACCGCTTGAAGGAGAGTTCAGAGGACCGCAGGACATCCGCCCAAC 180
DB 683 CTGTGCGACACCGCTTGAAGGAGAGTTCAGAGGACCGCAGGACATCCGCCCAAC 742
QY 181 ATCATCTGTGTCTTGAAGGAGCAGACAGATGTGGAGCTGGGTCCATGCAAGTGTGTAAC 240
DB 743 ATCATCTGTGTCTTGAAGGAGCAGACAGATGTGGAGCTGGGTCCATGCAAGTGTGTAAC 802
QY 241 AAGACCGCGCATCATGAGAGCAGGCGGGCGCATTCATCAACGCTTCGTGACCA 300
DB 803 AAGACCGCGCATCATGAGAGCAGGCGGGCGCATTCATCAACGCTTCGTGACCA 862
QY 301 CCCATGTGTCGCCCTCAGCTCTCCATCTCTCACTGCGCAAGTACGTCACACACCAAC 360
DB 863 CCCATGTGTCGCCCTCAGCTCTCCATCTCTCACTGCGCAAGTACGTCACACACCAAC 922
QY 361 ACCTACACCAATGAAATGTCTCTGCGCTCTTGGCAGGCACAGCAGAGCGCGC 420
DB 923 ACCTACACCAATGAAATGTCTCTGCGCTCTTGGCAGGCACAGCAGAGCGCGC 982
QY 421 ACCTTGGCGGTACCTCAATAGCACTGCTACCGGACAGCTTCTTCGGAGTATCT 480
DB 983 ACCTTGGCGGTACCTCAATAGCACTGCTACCGGACAGCTTCTTCGGAGTATCT 1042
QY 481 AATGAATACACGCGCTCTACGTGCCACCGGCTGGAGAGAGTGGGTGCGACTCTTAA 540
DB 1043 AATGAATACACGCGCTCTACGTGCCACCGGCTGGAGAGAGTGGGTGCGACTCTTAA 1102
QY 541 AACTCCCGCTTTTAACTACACGCTGTGTGGAACCGGGGTGAAGAAAGCAGCGCTCC 600
DB 1103 AACTCCCGCTTTTAACTACACGCTGTGTGGAACCGGGGTGAAGAAAGCAGCGCTCC 1162
QY 601 GACTACTCCAGGATTAACCTCACAGACTCATACCAATGACACGCTGAGCTTCTCCGC 660
DB 1163 GACTACTCCAGGATTAACCTCACAGACTCATACCAATGACACGCTGAGCTTCTCCGC 1222
QY 661 ACGTCCAGAGAGATGTATCCCGCAGAGCCAGTCTCATGTGTCATGACCATGAGCCCCC 720
DB 1223 ACGTCCAGAGAGATGTATCCCGCAGAGCCAGTCTCATGTGTCATGACCATGAGCCCCC 1282
QY 721 CACGCGCTGAGGATTCAGCCCAATATTCAGCGCTCTTCCCAACGCACTTCAGCAC 780
DB 1283 CACGCGCTGAGGATTCAGCCCAATATTCAGCGCTCTTCCCAACGCACTTCAGCAC 1342
QY 781 ATCAGCGGAGCTACAACTAGCGCCCAACCCCGGACAAACACTGGATCATGCGCTACAG 840
DB 1343 ATCAGCGGAGCTACAACTAGCGCCCAACCCCGGACAAACACTGGATCATGCGCTACAG 1402

QY 841 GGGCCCATGAAGCCCATCCATGAAATTCACCAATGCTCCAGCGAAGCGTTCAG 900
DB 1403 GGGCCCATGAAGCCCATCCATGAAATTCACCAATGCTCCAGCGAAGCGTTCAG 1462
QY 901 ACCCTCATGTGGTGGAGCACTCCATGAGAGCAATTTACACATGCTGTTGAGACGGC 960
DB 1463 ACCCTCATGTGGTGGAGCACTCCATGAGAGCAATTTACACATGCTGTTGAGACGGC 1522
QY 961 GAGCTGGAGCAACAGTATCATGCTATACACCGCCGACCAAGTTCACCATCGGCGAGTT 1020
DB 1523 GAGCTGGAGCAACAGTATCATGCTATACACCGCCGACCAAGTTCACCATCGGCGAGTT 1582
QY 1021 GGCCTGTGTAAGGGAATTCATGCGCATATGATTTGACATCAGGTTCCCTTCTACGTG 1080
DB 1583 GGCCTGTGTAAGGGAATTCATGCGCATATGATTTGACATCAGGTTCCCTTCTACGTG 1642
QY 1081 AGGGCCCCCAACAGTGGAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTCAGCTG 1140
DB 1643 AGGGCCCCCAACAGTGGAGCGGCTGTCTGAATCCCAACATCGTCTCAACATTCAGCTG 1702
QY 1141 GCGCCCAACATCTGGAACATTCAGGCGCTGGACATACCTGCGGATATGGAAGGAAATCC 1200
DB 1703 GCGCCCAACATCTGGAACATTCAGGCGCTGGACATACCTGCGGATATGGAAGGAAATCC 1762
QY 1201 ATCTCTAAGCTGTGGACACGAGCGCGGCTGTGAATCGGTTTCACTTGAAAGAGATG 1260
DB 1763 ATCTCTAAGCTGTGGACACGAGCGCGGCTGTGAATCGGTTTCACTTGAAAGAGATG 1822
QY 1261 AGGGTCTGCGCGGACTCTCTTCTGTGGAGAGAGCGAAGTGTCTACACAGAGAGACAAT 1320
DB 1823 AGGGTCTGCGCGGACTCTCTTCTGTGGAGAGAGCGAAGTGTCTACACAGAGAGACAAT 1882
QY 1321 GACAGGTGGAGCGCCAGGAGAGAACTTTCTGCGCAAGTACCAAGTGTGAAGACCTG 1380
DB 1883 GACAGGTGGAGCGCCAGGAGAGAACTTTCTGCGCAAGTACCAAGTGTGAAGACCTG 1942
QY 1381 TGTCAAGCTGTGTAGTACAGACGCGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440
DB 1943 TGTCAAGCTGTGTAGTACAGACGCGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 2002
QY 1441 GAGGACGCGACGCGGAGAGCTGAAGTGAAGTGAAGGCGCCCATGCGGCTGGCGGCGC 1500
DB 2003 GAGGACGCGACGCGGAGAGCTGAAGTGAAGTGAAGGCGCCCATGCGGCTGGCGGCGC 2062
QY 1501 AGCAGAGCCCTCTCCAACTGCTGCGCCAAAGTACTACGCGGAGGCGAGCGCTGCAAC 1560
DB 2063 AGCAGAGCCCTCTCCAACTGCTGCGCCAAAGTACTACGCGGAGGCGAGCGCTGCAAC 2122
QY 1561 TGTGACGCGGGGACTACAAAGCTCAGCTGCGCGGAGCGCGGAAAAAACTCTTCAAGAG 1620
DB 2123 TGTGACGCGGGGACTACAAAGCTCAGCTGCGCGGAGCGCGGAAAAAACTCTTCAAGAG 2182
QY 1621 AAGTACAGGCGCAGTATGTCGCGAGTGTCTCATCGCTCAGTGGCCATCGAGGTGAC 1680
DB 2183 AAGTACAGGCGCAGTATGTCGCGAGTGTCTCATCGCTCAGTGGCCATCGAGGTGAC 2242
QY 1681 GCGAGGTGTACCACTAGGCTGCGGTGATCGCGCCAGCGCCGAAAACTCACCAAGCGG 1740
DB 2243 GCGAGGTGTACCACTAGGCTGCGGTGATCGCGCCAGCGCCGAAAACTCACCAAGCGG 2302
QY 1741 CACTGCGCAGGCGCGCTGTGAGGACCAAGATGAAGGATGTGGGAGCTTCAGTGGGACT 1800
DB 2303 CACTGCGCAGGCGCGCTGTGAGGACCAAGATGAAGGATGTGGGAGCTTCAGTGGGACT 2362
QY 1801 GGAGGCTTCCGCACTACTCAGCGCGCAACCCCAATTAAGTGACACATCGGTGTACATC 1860
DB 2363 GGAGGCTTCCGCACTACTCAGCGCGCAACCCCAATTAAGTGACACATCGGTGTACATC 2422
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACTGAGCTGACCTGACAGTCCCTGCGAGCTGGAAA 1920
DB 2423 CTAGAGAACGACACAGTCCAGTGTGACTGAGCTGACAGTCCCTGCGAGCTGGAAA 2482
QY 1921 GACCACAGCTGCATCGACACGAGATTGAAACCTTGCAGAACAAATTAAGAACCTG 1980

Db 2483 GACCAAGCTGCATCGACACAGAGATTGAAACCTTCGAGAACAAATTAAGAACCTG 2542
QY 1981 AGGGAAGTCCGAGGTCACTGAAAGAAAGCGGCCAGAGAAATGTGACTGTGCACAAATC 2040
Db 2543 AGGGAAGTCCGAGGTCACTGAAAGAAAGCGGCCAGAGAAATGTGACTGTGCACAAATC 2602
QY 2041 AGCTACCAACCCAGCAGCAAAAGCGCCCTCAAGCAAGAGGCTCCAGTCTCATCTTTC 2100
Db 2603 AGCTACCAACCCAGCAGCAAAAGCGCCCTCAAGCAAGAGGCTCCAGTCTCATCTTTC 2662
QY 2101 AGGAAGGCTCCTCAAGAGAGAGCAAGGTGTGGCTGTTGCGGAGCAGAGCGCAAGAG 2160
Db 2663 AGGAAGGCTCCTCAAGAGAGAGCAAGGTGTGGCTGTTGCGGAGCAGAGCGCAAGAG 2722
QY 2161 AAATCTCGCAAGCTCTCAAGCGCTCGAGCAAAACGACAGCTGCAGATCCAGGCTC 2220
Db 2723 AAATCTCGCAAGCTCTCAAGCGCTCGAGCAAAACGACAGCTGCAGATCCAGGCTC 2782
QY 2221 AGTGTCTTCAACCCAGCAACACGACTGGCAGAGCGGCTTCTTGACACTGGGCT 2280
Db 2783 AGTGTCTTCAACCCAGCAACACGACTGGCAGAGCGGCTTCTTGACACTGGGCT 2842
QY 2281 TTCTGTGCTGCACAGCGCCCAACAACTACGTAAGTGTGATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGCACAGCGCCCAACAACTACGTAAGTGTGATGAGGACCATCAATGAG 2902
QY 2341 ACTCAAAATTCCTCTTCTGTGAATTTGCAACTGCTTCTAGAGTACTTTGATCTCAAC 2400
Db 2903 ACTCAAAATTCCTCTTCTGTGAATTTGCAACTGCTTCTAGAGTACTTTGATCTCAAC 2962
QY 2401 ACAGACCCCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
Db 2963 ACAGACCCCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3022
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 2520
Db 3023 CTACAGCTACAGCTCATGAGCTGAGGAGCTGCAAGGTTTCAAGCAGTGTAAACCCCGG 3082
QY 2521 ACTCGAATCATGAGCTGAGCTTAAAGATGAGGAGGATGATGATGATGATGATGATG 2580
Db 3083 ACTCGAATCATGAGCTGAGCTTAAAGATGAGGAGGATGATGATGATGATGATGATG 3133
QY 2581 CAGCGTCTGAAAGTGGCCAGAAATGAAGAGACCTTCTTCAAAATCACTGGGCAACTGTGG 2640
Db 3134 CAGCGTCTGAAAGTGGCCAGAAATGAAGAGACCTTCTTCAAAATCACTGGGCAACTGTGG 3193
QY 2641 GAAGCTGGAGGTTAAGAAACACAGAGGTGGACCTCCAAACACATAGAGGATCAC 2700
Db 3194 GAAGCTGGAGGTTAAGAAACACAGAGGTGGACCTCCAAACACATAGAGGATCAC 3253
QY 2701 TGAATGCACAGGCAATGAAACCACTGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760
Db 3254 TGAATGCACAGGCAATGAAACCACTGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 3313
QY 2761 AGAGGCTTGAGAAAGCAAGCAAGCACTCTAGTCAACATAGAGATTCCTGGAGATAAC 2820
Db 3314 AGAGGCTTGAGAAAGCAAGCAAGCACTCTAGTCAACATAGAGATTCCTGGAGATAAC 3373
QY 2821 CAGCAGGACAGAGATAACTTCAAGAGTCCATTTTTCCTGCTGCTTTTGTGTTGATTA 2880
Db 3374 CAGCAGGACAGAGATAACTTCAAGAGTCCATTTTTCCTGCTGCTTTTGTGTTGATTA 3433
QY 2881 TACCTTACAGCTGCAAAAATGCAATTTTTCCTGATCAAAAAGTCAACACTAACCTCCC 2940
Db 3434 TACCTTACAGCTGCAAAAATGCAATTTTTCCTGATCAAAAAGTCAACACTAACCTCCC 3493
QY 2941 CCAGAGCTCACAAGGAAAACGGAGAGAGCAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCACAAGGAAAACGGAGAGAGCAGAGAGATTTCTTGGAAATTTCT 3553
QY 3001 CCCAAGGCGGAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3060

Db 3554 CCCAAGGCGGAAGTCAATTTGAAATTTTAAATCATAGGGGAAAAGCAGTCTCTTTCTAAA 3613
QY 3061 TCTCTTATTTCTTTTGGTTTGTCAAAAGAGAACTAAGAGCAGGACAGAGGCAACGT 3120
Db 3614 TCTCTTATTTCTTTTGGTTTGTCAAAAGAGAACTAAGAGCAGGACAGAGGCAACGT 3673
QY 3121 GGAGAGGCTGAAACACAGTCAGAGACGTTTGAATAGTCAATAGTCAATAGGATGA 3180
Db 3674 GGAGAGGCTGAAACACAGTCAGAGACGTTTGAATAGTCAATAGTCAATAGGATGA 3733
QY 3181 CATTTACCTAGCACTATAAACCTGTGCTCTGAGAACTGCTTTCATTTGATATAT 3240
Db 3734 CATTTACCTAGCACTATAAACCTGTGCTCTGAGAACTGCTTTCATTTGATATAT 3793
QY 3241 GTGACTATTTTACATGTAATCAACATGAGGAACTTTTAGGGAACTTAATAAGAAATCCCAA 3300
Db 3794 GTGACTATTTTACATGTAATCAACATGAGGAACTTTTAGGGAACTTAATAAGAAATCCCAA 3853
QY 3301 TTTTTCAGGAGTGTGTGTCATTAACCGCTCTGTCGCCAGTGTAAAGAAAAA 3353
Db 3854 TTTTTCAGGAGTGTGTGTCATTAACCGCTCTGTCGCCAGTGTAAAGAAAAA 3906

RESULT 8
AX454488
LOCUS 3906 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 73 from Patent WO0208284.
ACCESSION AX454488
VERSION AX454488.1 GI:21713868
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Baker, K.P., Ferrara, N., Gerber, H., Gerritsen, M.E., Goddard, A.,
Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, J.,
Paoni, N.F., Stephan, J.P., Watanabe, C.K., Williams, P.M., Wood, W.I.
and Ye, W.
TITLE Compositions and methods for the diagnosis and treatment of
disorders involving angiogenesis
JOURNAL Patent: WO 0208284-A 73 31-JAN-2002;
Genentech, Inc. (US); Baker, Kevin P. (US); Ferrara, Napoleone
(US); Gerber, Hanspeter (US); Gerritsen, Mary E. (US); Goddard,
Audrey (US); Godowski, Paul J. (US); Gurney, Austin L. (US);
Hillan, Kenneth J. (US); Marsters, Scott A. (US); Pan, James (US);
Paoni, Nicholas F. (US); Stephan, Jean-Philippe F. (US);
Watanabe, Colin K. (US); Williams, P. Mickey (US); Wood, William
I. (US)

FEATURES
source
1..3906
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1047 a 1089 c 1041 g 729 t
ORIGIN
Query Match 98.8%; Score 3323.4; DB 6; Length 3906;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3343; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
QY 1 GACTCCGGATCCCAAGAGACAGATCAGCAAAAAGAGATGGCCCCCGAGC 60
Db 563 GACTCCGGATCCCAAGAGAGACAGATCAGCAAAAAGAGATGGCCCCCGAGC 622
QY 61 CTCGTGCTGTGCTGTGCTCCGCAACTGTGTCTCCCTGCTGGTGGAGCTCGGCTTC 120
Db 623 CTCGTGCTGTGCTGTGCTCCGCAACTGTGTCTCCCTGCTGGTGGAGCTCGGCTTC 682
QY 121 CTCGTGCAACCGCTGAAAGCGAGGTTTCAAGAGGACCGCAGGAAATCCGCCCCAC 180
Db 683 CTCGTGCAACCGCTGAAAGCGAGGTTTCAAGAGGACCGCAGGAAATCCGCCCCAC 742

181 ATATCTCTGCTGTGACGAGCAGCCAGAGTGTGGAGCTGGGTTCCATGCGAGTGTGAAC 240
743 ATATCTCTGCTGTGACGAGCAGCCAGAGTGTGGAGCTGGGTTCCATGCGAGTGTGAAC 802
241 AAGACCGGCGCATCATGAGACGAGGCGGCGGCGGACTTATCAACGCTTGTGACCA 300
803 AAGACCGGCGCATCATGAGACGAGGCGGCGGCGGACTTATCAACGCTTGTGACCA 862
301 CCATGTGCTGCGGCTTCAAGCTCTCTCATCTCTCACTGGCAAGTACGTCCACACCAAC 360
863 CCATGTGCTGCGGCTTCAAGCTCTCTCATCTCTCACTGGCAAGTACGTCCACACCAAC 922
361 ACCTACCAACCAATGAGACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
923 ACCTACCAACCAATGAGACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 982
421 ACCTTTCCGCTGTACCTCAATAGCACTGTGCTACCGGACAGCTTCTCTCGGAGTATCTT 480
983 ACCTTTCCGCTGTACCTCAATAGCACTGTGCTACCGGACAGCTTCTCTCGGAGTATCTT 1042
481 AATGAATPACCAACGCTCTCTACGTCGCCACCCGCTGGAAGAGTGGTTCGACTCTCTTAA 540
1043 AATGAATPACCAACGCTCTCTACGTCGCCACCCGCTGGAAGAGTGGTTCGACTCTCTTAA 1102
541 AACTCCGCTTTTATAACTACAGCTGTGTGCGAAAGGAGTGGTTCGACTCTCTTAA 600
1103 AACTCCGCTTTTATAACTACAGCTGTGTGCGAAAGGAGTGGTTCGACTCTCTTAA 1162
601 GACTACTTCCAGGATTAACCTCACAGACCTCATCAACCAATGACAGCGTGTCTTCCGC 660
1163 GACTACTTCCAGGATTAACCTCACAGACCTCATCAACCAATGACAGCGTGTCTTCCGC 1222
661 AGTTCGAAGAGTGTACCGGCAAGGCGAGTCTCATGTGTATGACGATGAGCGGCGG 720
1223 AGTTCGAAGAGTGTACCGGCAAGGCGAGTCTCATGTGTATGACGATGAGCGGCGG 1282
721 CAGGCGCTCAGGATTCAGCGCCACCAATATTCAGCGCTCTTCCAAACGCTCTCAGCAC 780
1283 CAGGCGCTCAGGATTCAGCGCCACCAATATTCAGCGCTCTTCCAAACGCTCTCAGCAC 1342
781 ATACGCGGAGCTACAACTACGCGGCCCAACCCGGAACAACTGTGATGCGCTACAG 840
1343 ATACGCGGAGCTACAACTACGCGGCCCAACCCGGAACAACTGTGATGCGCTACAG 1402
841 GGGCCCATGAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAGCGCTTGCAG 900
1403 GGGCCCATGAGCCCATCCACATGGAATTCACCAATGCTCCAGCGGAGCGCTTGCAG 1462
901 ACCCTCATGTGCTGAGGACTCCATGAGACGATTTACAACTGTGCTGAGACGGGC 960
1463 ACCCTCATGTGCTGAGGACTCCATGAGACGATTTACAACTGTGCTGAGACGGGC 1522
961 GAGTGTGACCAACAGTATCATGATATACCGCGGACCAAGTTACCACTGCGGCGAGTTT 1020
1523 GAGTGTGACCAACAGTATCATGATATACCGCGGACCAAGTTACCACTGCGGCGAGTTT 1582
1021 GGCCTGTGAAAGGAAATCCATGCTCATATGAGTTTGAATCAGGCTCCGCTTCTACG 1080
1583 GGCCTGTGAAAGGAAATCCATGCTCATATGAGTTTGAATCAGGCTCCGCTTCTACG 1642
1081 AGGGGCCCCAAGTGGAGCGGCTGTCTGAATCCCAATCCCAATCCCAATCCCAATCCCA 1140
1643 AGGGGCCCCAAGTGGAGCGGCTGTCTGAATCCCAATCCCAATCCCAATCCCAATCCCA 1702
1141 GCGCCACCATCTGAGCATTTGAGGCGCTGACATACCTGCGATATGAGCGGGAATCC 1200
1703 GCGCCACCATCTGAGCATTTGAGGCGCTGACATACCTGCGATATGAGCGGGAATCC 1762
1201 ATCTCAAGCTGTGACAGCGGCGGCGGCTGAATCGGTTTCACTTGAAGAAAGATG 1260
1763 ATCTCAAGCTGTGACAGCGGCGGCGGCTGAATCGGTTTCACTTGAAGAAAGATG 1822
1261 AGGCTCTGGCGGACTCTCTTCTGCTGGAGAGGAGGCTGTCTACACAGAGACAAAT 1320

1823 AGGCTCTGGCGGACTCTCTTCTGCTGAGAGAGGCAAGCTGTCTACACAGAGAGCAAT 1882
1321 GACAAAGTGTGACGCGCCAGGAGAGAACTTTCTGCCCAAGTACCAAGCTGTGAAGAGCTG 1380
1883 GACAAAGTGTGACGCGCCAGGAGAGAACTTTCTGCCCAAGTACCAAGCTGTGAAGAGCTG 1942
1381 TGTCAAGCTGTGAGTATACCAAGACGCGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 1440
1943 TGTCAAGCTGTGAGTATACCAAGACGCGCTGTGAGCAGCTGGGACAGAAAGTGGCAGTGTGTG 2002
1441 GAGGACGCCACCGGGAAGCTGAAGCTGCATTAAGTGAAGGCGCCCATGCGGCTGGGCGGC 1500
2003 GAGGACGCCACCGGGAAGCTGAAGCTGCATTAAGTGAAGGCGCCCATGCGGCTGGGCGGC 2062
1501 AGCAGAGCCCTCTCCAACTCTGCTCCCAAGTACTACGCGGAGGCGGACGAGGCGCTGCACC 1560
2063 AGCAGAGCCCTCTCCAACTCTGCTCCCAAGTACTACGCGGAGGCGGACGAGGCGCTGCACC 2122
1561 TGTGACGCGGGAAGTACAAAGCTCAGCTGCGCGGACGCGGGAAGAACTCTTCAAGAG 1620
2123 TGTGACGCGGGAAGTACAAAGCTCAGCTGCGCGGACGCGGGAAGAACTCTTCAAGAG 2182
1621 AAGTACAGGCGCAGCTATGTCGAGTCTGCTCCATCCGCTCAGTGGCCATCGAGCTGAC 1680
2183 AAGTACAGGCGCAGCTATGTCGAGTCTGCTCCATCCGCTCAGTGGCCATCGAGCTGAC 2242
1681 GGCAGGCTGTACCAAGCTAGGCTGCGGCTGATGCGCGCCAGCGCCCGAAACCTCACCAGCGG 1740
2243 GGCAGGCTGTACCAAGCTAGGCTGCGGCTGATGCGCGCCAGCGCCCGAAACCTCACCAGCGG 2302
1741 CACTGGCCAGGCGCCCTGAGGACCAAGATGACAGGATGCTGGGAGCTTCAGTGGCACT 1800
2303 CACTGGCCAGGCGCCCTGAGGACCAAGATGACAGGATGCTGGGAGCTTCAGTGGCACT 2362
1801 GZAGGCTTCCGCACTACTCAGCGCCCAACCCCAATTAAGTGAACATCGGTGTACATC 1860
2363 GZAGGCTTCCGCACTACTCAGCGCCCAACCCCAATTAAGTGAACATCGGTGTACATC 2422
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAGTCCCTGCGAGGCTGAAA 1920
2423 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAGTCCCTGCGAGGCTGAAA 2482
1921 GACCAGAGCTGACATCGACACGAGATTGAACCTCTGAGAACCAAAATTAAGAACTG 1980
2483 GACCAGAGCTGACATCGACACGAGATTGAACCTCTGAGAACCAAAATTAAGAACTG 2542
1981 AGGGAAGTCCGAGTCTACCTGAAGAAAAGCGGCCAGAGAAATGTGACTGTCAAAAATC 2040
2543 AGGGAAGTCCGAGTCTACCTGAAGAAAAGCGGCCAGAGAAATGTGACTGTCAAAAATC 2602
2041 AGCTACCAACCCAGACCAAGGCGGCTCAGGACAGAGGCTCCAGTCTGCACTCTTTC 2100
2603 AGCTACCAACCCAGACCAAGGCGGCTCAGGACAGAGGCTCCAGTCTGCACTCTTTC 2662
2101 AGGAAGGCTTGCAGAGAGAGGACAAAGTGTGGCTGTGGGAGGAGGAGCGCAAGAG 2160
2663 AGGAAGGCTTGCAGAGAGAGGACAAAGTGTGGCTGTGGGAGGAGGAGCGCAAGAG 2722
2161 AAATCCGCAAGCTGTCTAAGCGGCTCGAGAACAAAGACGCTGACGATGCGAGGCTC 2220
2723 AAATCCGCAAGCTGTCTAAGCGGCTCGAGAACAAAGACGCTGACGATGCGAGGCTC 2782
2221 ACGTGCTTCCACACGACCAACAGCACTGGCAGAGCGGCTTCTGAGACACTGGGCGCT 2280
2783 ACGTGTCTTCCACACGACCAACAGCACTGGCAGAGCGGCTTCTGAGACACTGGGCGCT 2842
2281 TTCTGTGCTGTGACACGAGCGCAACAAATAACAGTACTGGTGTGATGAGGACCATCAATGAG 2340
2843 TTCTGTGCTGTGACACGAGCGCAACAAATAACAGTACTGGTGTGATGAGGACCATCAATGAG 2902
2341 ACTCACAATTTCTTCTGTGTGATTTGCACTTCTAGAGTACTTGTGATCTCAAC 2400

Db 2903 ACTCACAATTTCTCTTCTGTGAATTTTGCAACTGGCTTCTAGAGTACTTTGTATCTCAAC 2962
QY 2401 ACAGACCCCTACAGCTGATGAATGTCAGTGAACACACTGGACAGGATGCTCTCAACACAG 2460
Db 2963 ACAGACCCCTACAGCTGATGAATGTCAGTGAACACACTGGACAGGATGCTCTCAACACAG 3022
QY 2461 CTACACGTACAGCTATGGAGCTGAGGAGTGCAGAGGTTTACAGCAGTGTAAACCCCGG 2520
Db 3023 CTACACGTACAGCTATGGAGCTGAGGAGTGCAGAGGTTTACAGCAGTGTAAACCCCGG 3082
QY 2521 ACTCGAAACATGGAGCTGGAGCTTAAAGATGAGGAGGAGCTATGAGCAATACAGCAGTTT 2580
Db 3083 ACTCGAAACATGGAGCTGGAGCTTAAAGATGAGGAGGAGCTATGAGCAATACAGCAGTTT 3133
QY 2581 CAGCGTCGAAGTGCCAGAAATGAAGAGACTCTTCCAAATCACTGGACAACTGTGG 2640
Db 3134 CAGCGTCGAAGTGCCAGAAATGAAGAGACTCTTCCAAATCACTGGACAACTGTGG 3193
QY 2641 GAAGCTGGAGGTTTAAAGAACACAGAGTGGAGCTCCAAACATAGAGCCTCAAC 2700
Db 3194 GAAGCTGGAGGTTTAAAGAACACAGAGTGGAGCTCCAAACATAGAGCCTCAAC 3253
QY 2701 TGACTGCAAGCAATGAAACCACTGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 2760
Db 3254 TGACTGCAAGCAATGAAACCACTGTGGTGATTTCCAGCAGACCTGTGCTATTGGCC 3313
QY 2761 AGGAGCTGGAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATCTGGAGGATAAC 2820
Db 3314 AGGAGCTGGAAGCAAGCAGCAGCTCTCAGTCAACATGACAGATCTGGAGGATAAC 3373
QY 2821 CAGCAGGAGCAGAGATACTTTCAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 2880
Db 3374 CAGCAGGAGCAGAGATACTTTCAGAGTCCATTTTGGCCCTGCTTTGCTTTGGATTA 3433
QY 2881 TACCTCAACAGCTGCAAAAATGCAATTTTTCGATCAAAAAGTCAACATAACCTCTCC 2940
Db 3434 TACCTCAACAGCTGCAAAAATGCAATTTTTCGATCAAAAAGTCAACATAACCTCTCC 3493
QY 2941 CCAGAGCTCAAGAGGAAACGAGAGAGCAGCAGAGAGATTTCTTGGAAATTTCT 3000
Db 3494 CCAGAGCTCAAGAGGAAACGAGAGAGCAGCAGAGAGATTTCTTGGAAATTTCT 3553
QY 3001 CCCAAGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3060
Db 3554 CCCAAGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAGCAGTCTGTCTTAAA 3613
QY 3061 TCCTCTTATTTCTTTGTTGTCACAAAGAGGAACTTAAGAGCAGCAGAGGCAACGT 3120
Db 3614 TCCTCTTATTTCTTTGTTGTCACAAAGAGGAACTTAAGAGCAGCAGAGGCAACGT 3673
QY 3121 GGAGAGGCTGAAACAGTGCAGAGACGTTTGACATGAGTCAGTAGCAGCAAAAAGAGATGA 3180
Db 3674 GGAGAGGCTGAAACAGTGCAGAGACGTTTGACATGAGTCAGTAGCAGCAAAAAGAGATGA 3733
QY 3181 CATTTACCTAGCATAAACCCTGGTTCCTCTGAAGAACTCGCTTCATGTATATAT 3240
Db 3734 CATTTACCTAGCATAAACCCTGGTTCCTCTGAAGAACTCGCTTCATGTATATAT 3793
QY 3241 GTGACTATTATACATGTAACATGGAACCTTTTGGGGAACTTAAGAGAAATCCCAA 3300
Db 3794 GTGACTATTATACATGTAACATGGAACCTTTTGGGGAACTTAAGAGAAATCCCAA 3853
QY 3301 TTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA 3353
Db 3854 TTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA 3906

RESULT 9
AX463886
LOCUS AX463886 3906 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 19 from Patent WO0140466.
ACCESSION AX463886
VERSION AX463886.1 GI:21898960

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E., Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L., Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K., Wood, W.L. and Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0140466-A 19 07-JUN-2001;
Genentech Inc. (US)
FEATURES
Location/Qualifiers
source 1..3906
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1047 a 1089 c 1041 g 729 t
ORIGIN
Query Match 98.8%; Score 3323.4; DB 6; Length 3906;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3343; Conservative 1; Mismatches 9; Gaps 1;
QY 1 GACTCCCGATCCCAAAAGAGCACCAGATCAGCAAAAAGAGATGGCCCCCGAGC 60
Db 563 GACTCCCGATCCCAAAAGAGCACCAGATCAGCAAAAAGAGATGGCCCCCGAGC 622
QY 61 CTGCTGCTGTGCTGCTGCTGCGCACTGTGTTCTCCCTGCTGGGTGGAGCTCGGCCCTTC 120
Db 623 CTGCTGCTGTGCTGCTGCTGCGCACTGTGTTCTCCCTGCTGGGTGGAGCTCGGCCCTTC 682
QY 121 CTGCTGCAACCGCTCGAAAGCAGGTTTCAGAGGACCGCAGAGAAATCGCCCCCAAC 180
Db 683 CTGCTGCAACCGCTCGAAAGCAGGTTTCAGAGGACCGCAGAGAAATCGCCCCCAAC 742
QY 181 ATCATCTGCTGCTCTACAGCAGCAGCAGATGTGGAGCTGGGTTCATCGAGGTGATGAAC 240
Db 743 ATCATCTGCTGCTCTACAGCAGCAGCAGATGTGGAGCTGGGTTCATCGAGGTGATGAAC 802
QY 241 AAGACCCCGCGCATCATGAGCAGGCGGGGCGCACTTCATCAACGCTTCGTGACCA 300
Db 803 AAGACCCCGCGCATCATGAGCAGGCGGGGCGCACTTCATCAACGCTTCGTGACCA 862
QY 301 CCCATGTGCTGCCCTCAGCTCCCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 863 CCCATGTGCTGCCCTCAGCTCCCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 922
QY 361 ACCTACACCAACAATGAGAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 923 ACCTACACCAACAATGAGAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 982
QY 421 ACCTTTGCGGTGTACTCTCAATAGCACTGGCTACCGACAGCTTTCTTCGGGAGATATCTT 480
Db 983 ACCTTTGCGGTGTACTCTCAATAGCACTGGCTACCGACAGCTTTCTTCGGGAGATATCTT 1042
QY 481 AATGAATACAAACGGCTCTCTACGTGCCACCCGCTCGAAGAGAGTGGGTGGGTCTCTTAAA 540
Db 1043 AATGAATACAAACGGCTCTCTACGTGCCACCCGCTCGAAGAGAGTGGGTGGGTCTCTTAAA 1102
QY 541 ACTTCCCGCTTTTATTAATTAACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 1103 ACTTCCCGCTTTTATTAACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1162
QY 601 GACTACTTCAAGGATTTACCTCAGACAGCTCATCAACATGACAGCTGAGCTTCTTCGCG 660
Db 1163 GACTACTTCAAGGATTTACCTCAGACAGCTCATCAACATGACAGCTGAGCTTCTTCGCG 1222
QY 661 AGTCCAAAGAGATGTACCCGACAGGCGAGTCTCTCATGTGTATCATGAGCATGACGCCCC 720
Db 1223 AGTCCAAAGAGATGTACCCGACAGGCGAGTCTCTCATGTGTATCATGAGCATGACGCCCC 1282

QY 721 CACGGCCCTGAGGATTACAGCCCAACATATTCAGGCTCTTCCCAAAACGATCTCAGCAC 780
Db 1283 CACGGCCCTGAGGATTACAGCCCAACATATTCAGGCTCTTCCCAAAACGATCTCAGCAC 1342
QY 781 ATCAGCGCGAGCTCAACTAGCGGCCCAACCGGACAAACACTGGATCATCGCGTACAG 840
Db 1343 ATCAGCGCGAGCTCAACTAGCGGCCCAACCGGACAAACACTGGATCATCGCGTACAG 1402
QY 841 GGGCCCATGAAGCCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTCAG 900
Db 1403 GGGCCCATGAAGCCCATCCACATGGAAATTCACCAACATGCTCCAGCGGAAGCGCTTCAG 1462
QY 901 ACCCTCATGTGCGTGAGAGACTCATGAGAGAGATTTACAACTGCTGGTTGAGAGCGGC 960
Db 1463 ACCCTCATGTGCGTGAGAGACTCATGAGAGAGATTTACAACTGCTGGTTGAGAGCGGC 1522
QY 961 GAGCTGGACAAACACTCATCGTATACACCGCCGACCAAGCTTACCAACATCGCGCAGTTT 1020
Db 1523 GAGCTGGACAAACACTCATCGTATACACCGCCGACCAAGCTTACCAACATCGCGCAGTTT 1582
QY 1021 GGCTGTGTGAAGGAAATCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG 1080
Db 1583 GGCTGTGTGAAGGAAATCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTG 1642
QY 1081 AGGGGCCCCPACGTGGAGCGGCTGTGTAATCCCAACATCGTCTCAACATGACCTG 1140
Db 1643 AGGGGCCCCPACGTGGAGCGGCTGTGTAATCCCAACATCGTCTCAACATGACCTG 1702
QY 1141 GCCCCACCATCTGTGACATTTGACGGCTTGACATCTCGGATATGGACGGGAAATCC 1200
Db 1703 GCCCCACCATCTGTGACATTTGACGGCTTGACATCTCGGATATGGACGGGAAATCC 1762
QY 1201 ATCTCTAAGCTGTGACACAGGAGCGCGGTGTAATCGGTTTCACTTGAAGAAAGAGATG 1260
Db 1763 ATCTCTAAGCTGTGACACAGGAGCGCGGTGTAATCGGTTTCACTTGAAGAAAGAGATG 1822
QY 1261 AGGCTCTGGCGGACTCCTCTTGGTGAGAGAGCGCAAGCTGTACACAGAGAGACAAT 1320
Db 1823 AGGCTCTGGCGGACTCCTCTTGGTGAGAGAGCGCAAGCTGTACACAGAGAGACAAT 1882
QY 1321 GACAAAGTGTGACCGCCAGGAGGAACTTTCTGCCAAGTACCAAGCTGTGAAGGACCTG 1380
Db 1883 GACAAAGTGTGACCGCCAGGAGGAACTTTCTGCCAAGTACCAAGCTGTGAAGGACCTG 1942
QY 1381 TGTGAGCGTGTGAGTACCAAGCGCGCTGTGAGAGCTGGGACGAGTGGCAGTGTG 1440
Db 1943 TGTGAGCGTGTGAGTACCAAGCGCGCTGTGAGAGCTGGGACGAGTGGCAGTGTG 2002
QY 1441 GAGGACGCGACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCCATGCGGCTGGGCGGC 1500
Db 2003 GAGGACGCGACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCCATGCGGCTGGGCGGC 2062
QY 1501 AGCAGAGCCCTCTCAACCTGTGCCCAAGTACTACGGGACGCGGACGAGGCGCTGAC 1560
Db 2063 AGCAGAGCCCTCTCAACCTGTGCCCAAGTACTACGGGACGCGGACGAGGCGCTGAC 2122
QY 1561 TGTGACAGCGGGGACTACAAGCTCAGCGCTGCGCGAGCGCGGAAATACTCTTCAAGAAG 1620
Db 2123 TGTGACAGCGGGGACTACAAGCTCAGCGCTGCGCGAGCGCGGAAATACTCTTCAAGAAG 2182
QY 1621 AAGTACAAGGCGACGTAATGTCGCGAGTGGTCTCATCGCTCAGTGGCCATTCGAGGTGGAC 1680
Db 2183 AAGTACAAGGCGACGTAATGTCGCGAGTGGTCTCATCGCTCAGTGGCCATTCGAGGTGGAC 2242
QY 1681 GGCAGGTTGTACACGATGAGCGCTGGGTGATGTCGCCCGCCAGCGCCGAACTCACCAGAGCG 1740
Db 2243 GGCAGGTTGTACACGATGAGCGCTGGGTGATGTCGCCCGCCAGCGCCGAACTCACCAGAGCG 2302
QY 1741 CACTGGCCAGGGGCCCTGTGAGGACCAAGATGACAAGGATGGTGGGACTTTCAGTGGCACT 1800
Db 2303 CACTGGCCAGGGGCCCTGTGAGGACCAAGATGACAAGGATGGTGGGACTTTCAGTGGCACT 2362

QY 1801 GGAGGCTTCCCGACTACTCAGCGCCCAACCCCAATTAAGTGACACATCGGTGCTACATC 1860
Db 2363 GGAGGCTTCCCGACTACTCAGCGCCCAACCCCAATTAAGTGACACATCGGTGCTACATC 2422
QY 1861 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTGTAAGTCCCTGAGGCTTGAAAA 1920
Db 2423 CTAGAGAACGACACAGTCCAGTGTGACCTGACCTGTGTAAGTCCCTGAGGCTTGAAAA 2482
QY 1921 GACCAACAGCTGCACATCGACCACAGAGATTGAAACCTCTGCAGAACAAAAATTAAAGAACCTG 1980
Db 2483 GACCAACAGCTGCACATCGACCACAGAGATTGAAACCTCTGCAGAACAAAAATTAAAGAACCTG 2542
QY 1981 AGGGAAGTCCGAGGTTCACCTGAAGAAAAAGCGGCCAGAGAAATGTGATGTCACAAAAATC 2040
Db 2543 AGGGAAGTCCGAGGTTCACCTGAAGAAAAAGCGGCCAGAGAAATGTGATGTCACAAAAATC 2602
QY 2041 AGCTACCAACCCACGACAAAGAGCGGCTCTCAAGACACAGAGCTCCAGTCTGATCTTTC 2100
Db 2603 AGCTACCAACCCACGACAAAGAGCGGCTCTCAAGACACAGAGCTCCAGTCTGATCTTTC 2662
QY 2101 AGGAAGGCGCTCAAGAGAGGACAAAGGTGTGGCTGTTCGGGAGCAGAAAGCGCAAGAG 2160
Db 2663 AGGAAGGCGCTCAAGAGAGGACAAAGGTGTGGCTGTTCGGGAGCAGAAAGCGCAAGAG 2722
QY 2161 AAATCTCCGAGCTGCTCAAGCGGCTGACAAACACACACACAGTGCAGATGCCAGGCTC 2220
Db 2723 AAATCTCCGAGCTGCTCAAGCGGCTGACAAACACACACACAGTGCAGATGCCAGGCTC 2782
QY 2221 AGTGTCTTCCACCCAGACAAACAGCACTGGGAGAGCGGCGCTTCTGGACACTGGGGCT 2280
Db 2783 AGTGTCTTCCACCCAGACAAACAGCACTGGGAGAGCGGCGCTTCTGGACACTGGGGCT 2842
QY 2281 TTCTGTGCTGACACAGCGCCAAATTAACACGTACTGTGTGATGAGGACCATCAATGAG 2340
Db 2843 TTCTGTGCTGACACAGCGCCAAATTAACACGTACTGTGTGATGAGGACCATCAATGAG 2902
QY 2341 ACTCAAAATTTCTTCTTGTGAAATTTGCAACTTGTCTTCTAGAGTACTTTGATCTCAAC 2400
Db 2903 ACTCAAAATTTCTTCTTGTGAAATTTGCAACTTGTCTTCTAGAGTACTTTGATCTCAAC 2962
QY 2401 ACAGACCCCTACAGCTGATGAATGACAGTGAACACACTGGACAGGGATGTCTCAACCCAG 2460
Db 2963 ACAGACCCCTACAGCTGATGAATGACAGTGAACACACTGGACAGGGATGTCTCAACCCAG 3022
QY 2461 CTACAGCTACAGCTCATGAGCTGAGGCTGCAAGGGTTACAGAGTGTAAACCCCGG 2520
Db 3023 CTACAGCTACAGCTCATGAGCTGAGGCTGCAAGGGTTACAGAGTGTAAACCCCGG 3082
QY 2521 ACTCAAAACATGGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGGCACTTT 2580
Db 3083 ACTCAAAACATGGACCTG-----GATGGAGGAAGCTATGAGCAATACAGGCACTTT 3133
QY 2581 CAGCGTCAAAAGTGGCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 2640
Db 3134 CAGCGTCAAAAGTGGCAGAAATGAAGAGACTTCTTCCAAATCACTGGGACACTGTGG 3193
QY 2641 GAAGGCTGGGAAGGTTAAGAAACAAACAGAGGTGACCTCCAAAAACATAGAGGATCACC 2700
Db 3194 GAAGGCTGGGAAGGTTAAGAAACAAACAGAGGTGACCTCCAAAAACATAGAGGATCACC 3253
QY 2701 TGACTGCAACGGCAATGAAGAAACCATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 2760
Db 3254 TGACTGCAACGGCAATGAAGAAACCATGTGGGTGATTTCCAGCAGACCTGTGTATTGGCC 3313
QY 2761 AGGAGGCTGGGAAGGAAAGCAAGCAAGCTCTCAGTCAACATGACAGATTTCTGGAGATAAC 2820
Db 3314 AGGAGGCTGGGAAGGAAAGCAAGCAAGCTCTCAGTCAACATGACAGATTTCTGGAGATAAC 3373
QY 2821 CAGCAGGAGCAGAGATAACTTTTCAGAAAGTCCATTTTGGCCCTGCTTTTGTGTTGATTA 2880
Db 3374 CAGCAGGAGCAGAGATAACTTTTCAGAAAGTCCATTTTGGCCCTGCTTTTGTGTTGATTA 3433
QY 2881 TACTCACCAGCTGCACAAAAATGATTTTTTCGTATCAAAAAAGTCAACCACTAACCCCTCCC 2940

1763	ATCCTCAAGCTGCTGGACACAGGACGGCCGGTGAATCCGTTTCATCTGAAAAAGAGATG	1822
QY	AGGGTCTGGCGGACTCTCTTCTTGGTGAGAGAGCAAGCTGTCTACACAAGAGAGACAAT	1320
	AGGGTCTGGCGGACTCTCTTCTTGGTGAGAGAGCAAGCTGTCTACACAAGAGAGACAAT	1882
	AGGGTCTGGCGGACTCTCTTCTTGGTGAGAGAGCAAGCTGTCTACACAAGAGAGACAAT	1882
QY	GACAAGGTGGACCCCAAGGAGGAGAACTTCTCCCAAGTACACAGCGTGTGAAGGACCTG	1380
	GACAAGGTGGACCCCAAGGAGGAGAACTTCTCCCAAGTACACAGCGTGTGAAGGACCTG	1942
	GACAAGGTGGACCCCAAGGAGGAGAACTTCTCCCAAGTACACAGCGTGTGAAGGACCTG	1440
QY	TGTCAGCGTGTGAGTACCAAGACGGCGTGTGACAGCTGGGACAGAAAGTGGCAAGTGTGTG	1440
	TGTCAGCGTGTGAGTACCAAGACGGCGTGTGACAGCTGGGACAGAAAGTGGCAAGTGTGTG	2002
	TGTCAGCGTGTGAGTACCAAGACGGCGTGTGACAGCTGGGACAGAAAGTGGCAAGTGTGTG	2002
QY	GAGGACGCCACGGGGAGCTGAAGCTGCATAAGTGCAAAGGCCCCCATCGGGCTGGCGCGC	1500
	GAGGACGCCACGGGGAGCTGAAGCTGCATAAGTGCAAAGGCCCCCATCGGGCTGGCGCGC	2062
	GAGGACGCCACGGGGAGCTGAAGCTGCATAAGTGCAAAGGCCCCCATCGGGCTGGCGCGC	2062
QY	AGCAGAGCCCTCTCCAACTCTGTGCCCAAGTACTAGCGGACAGGACGAGGCGCTGCACC	1560
	AGCAGAGCCCTCTCCAACTCTGTGCCCAAGTACTAGCGGACAGGACGAGGCGCTGCACC	2122
	AGCAGAGCCCTCTCCAACTCTGTGCCCAAGTACTAGCGGACAGGACGAGGCGCTGCACC	2122
QY	TGTGA CAGCGGGGACTACAAAGCTCAGCTTGCGCGGACGCGCGGAAAAAACTCTTTCAAGAAG	1620
	TGTGA CAGCGGGGACTACAAAGCTCAGCTTGCGCGGACGCGCGGAAAAAACTCTTTCAAGAAG	2182
	TGTGA CAGCGGGGACTACAAAGCTCAGCTTGCGCGGACGCGCGGAAAAAACTCTTTCAAGAAG	2182
QY	AAGTACAAGGCCAGCTATGTCCGAGTTCGCTCCATCCGCTCAGTGGCGCATCGAGGTGGAC	1680
	AAGTACAAGGCCAGCTATGTCCGAGTTCGCTCCATCCGCTCAGTGGCGCATCGAGGTGGAC	2242
	AAGTACAAGGCCAGCTATGTCCGAGTTCGCTCCATCCGCTCAGTGGCGCATCGAGGTGGAC	2242
QY	GGCAGGGTGTACCACTGATAGGCTTGGGTGATGTCGCGCCAGCCCGCGAAACCTTCACCAAGCGG	1740
	GGCAGGGTGTACCACTGATAGGCTTGGGTGATGTCGCGCCAGCCCGCGAAACCTTCACCAAGCGG	2302
	GGCAGGGTGTACCACTGATAGGCTTGGGTGATGTCGCGCCAGCCCGCGAAACCTTCACCAAGCGG	2302
QY	CAGTGGCCAGGGGCCCTTGAGGACCAAGATGACAAGATGGTGGGACNTTCAGTGGCACT	1800
	CAGTGGCCAGGGGCCCTTGAGGACCAAGATGACAAGATGGTGGGACNTTCAGTGGCACT	2362
	CAGTGGCCAGGGGCCCTTGAGGACCAAGATGACAAGATGGTGGGACNTTCAGTGGCACT	2362
QY	GGAGGGCTTCTCCGACTACTCAGCGCGCAACCCCATTTAAAGTGCACATCGGTGCTACATC	1860
	GGAGGGCTTCTCCGACTACTCAGCGCGCAACCCCATTTAAAGTGCACATCGGTGCTACATC	2422
	GGAGGGCTTCTCCGACTACTCAGCGCGCAACCCCATTTAAAGTGCACATCGGTGCTACATC	2422
QY	CTAGAGAACGACACAGTCTGACTGTGACCTGGACCTGTGA CAAAGTCCCTGCAGGCTCGAAA	1920
	CTAGAGAACGACACAGTCTGACTGTGACCTGGACCTGTGA CAAAGTCCCTGCAGGCTCGAAA	2482
	CTAGAGAACGACACAGTCTGACTGTGACCTGGACCTGTGA CAAAGTCCCTGCAGGCTCGAAA	2482
QY	GACCAACAAGCTGCACATCGACACAGAGATTGAACCCCTGCAGAAACAAAATTAAAGAACCTG	1980
	GACCAACAAGCTGCACATCGACACAGAGATTGAACCCCTGCAGAAACAAAATTAAAGAACCTG	2542
	GACCAACAAGCTGCACATCGACACAGAGATTGAACCCCTGCAGAAACAAAATTAAAGAACCTG	2542
QY	AGGGAAGTCCGAGGTCACTGTGAAGAAAAAGCGGCCAGAGAAATGTGACTGTCAAAAAATC	2040
	AGGGAAGTCCGAGGTCACTGTGAAGAAAAAGCGGCCAGAGAAATGTGACTGTCAAAAAATC	2602
	AGGGAAGTCCGAGGTCACTGTGAAGAAAAAGCGGCCAGAGAAATGTGACTGTCAAAAAATC	2602
QY	AGCTACACACCCAGACAAAGCGCCCTCAAGCACAAGAGGCTCCAGTCTGCACTCTTTTC	2100
	AGCTACACACCCAGACAAAGCGCCCTCAAGCACAAGAGGCTCCAGTCTGCACTCTTTTC	2662
	AGCTACACACCCAGACAAAGCGCCCTCAAGCACAAGAGGCTCCAGTCTGCACTCTTTTC	2662
QY	AGGAAGGGCTCGAAGAAAGGACAAGGTGTGGTGTGGGGAGCAGAAAGCGCAAGAAG	2160
	AGGAAGGGCTCGAAGAAAGGACAAGGTGTGGTGTGGGGAGCAGAAAGCGCAAGAAG	2722
	AGGAAGGGCTCGAAGAAAGGACAAGGTGTGGTGTGGGGAGCAGAAAGCGCAAGAAG	2722
QY	AAACTCGGAAAGCTGTCTAAGCGCTTCAGAAACAA CAGACAGTGTGACGATCGCAGGCTC	2220
	AAACTCGGAAAGCTGTCTAAGCGCTTCAGAAACAA CAGACAGTGTGACGATCGCAGGCTC	2782
	AAACTCGGAAAGCTGTCTAAGCGCTTCAGAAACAA CAGACAGTGTGACGATCGCAGGCTC	2782
QY	ACGTGTCTTACCACAACACAGCACTGGCAGACGGCGCTTTCTGGAACA CTGGGGCCT	2280
	ACGTGTCTTACCACAACACAGCACTGGCAGACGGCGCTTTCTGGAACA CTGGGGCCT	2842
	ACGTGTCTTACCACAACACAGCACTGGCAGACGGCGCTTTCTGGAACA CTGGGGCCT	2842
QY	TTCTGTGCTTCGACACGAGCGCAACAATAACAGCTACTGGTGCATGAGGACCATCAATGAG	2340
	TTCTGTGCTTCGACACGAGCGCAACAATAACAGCTACTGGTGCATGAGGACCATCAATGAG	2902
	TTCTGTGCTTCGACACGAGCGCAACAATAACAGCTACTGGTGCATGAGGACCATCAATGAG	2902

QY	2341	ACTCAAAATTTCTCTTCGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC	2400
Db	2903	ACTCAAAATTTCTCTTCGTGAAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCAAC	2962
QY	2401	ACAGACCCCTACAGCTGATGAATGCAGTGAACACACTGCGACAGGGATGTCCTCAACCAG	2460
Db	2963	ACAGACCCCTACAGCTGATGAATGCAGTGAACACACTGCGACAGGGATGTCCTCAACCAG	3022
QY	2461	CTACAGTACAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGTAAACCCCGG	2520
Db	3023	CTACAGTACAGCTCATGAGCTGAGGAGCTCAAGGGTTACAAGCAGTGTAAACCCCGG	3082
QY	2521	ACTCGAATCATGACCTGGGACTTAAAGATGAGGAGAGCTATGAGCAATACAGGCAGTTT	2580
Db	3083	ACTCGAATCATGACCTGGGACTTAAAGATGAGGAGAGCTATGAGCAATACAGGCAGTTT	3133
QY	2581	CAGCGTCGAAAGTGGCCGAAATGAAGAGACCTTCTTCCAAATCACTGGGCAACTGTGG	2640
Db	3134	CAGCGTCGAAAGTGGCCGAAATGAAGAGACCTTCTTCCAAATCACTGGGCAACTGTGG	3193
QY	2641	GAAGGCTGGGAAGGTTAAGAAACACAGAGGTGGACCTTCCAAAATCACTGGGCAACTGTGG	2700
Db	3194	GAAGGCTGGGAAGGTTAAGAAACACAGAGGTGGACCTTCCAAAATCACTGGGCAACTGTGG	3253
QY	2701	TGACTGCACAGCAATGAAAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC	2760
Db	3254	TGACTGCACAGCAATGAAAAACCAATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCC	3313
QY	2761	AGAGGCTTGAGAAAGCAAGCACCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC	2820
Db	3314	AGAGGCTTGAGAAAGCAAGCACCTCTCAGTCAACATGACAGATTCCTGGAGGATAAC	3373
QY	2821	CAGCAGGACAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTAA	2880
Db	3374	CAGCAGGACAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTGCTTTGGATTAA	3433
QY	2881	TACCTCACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCACCACTAACCTCC	2940
Db	3434	TACCTCACAGCTGCACAAAATGCAATTTTTCGTATCAAAAAGTCACCACTAACCTCC	3493
QY	2941	CCAGAGCTCACAAAGGAAAACGGAGAGCGAGAGAGATTTCTTGGAAATTTCT	3000
Db	3494	CCAGAGCTCACAAAGGAAAACGGAGAGCGAGAGAGATTTCTTGGAAATTTCT	3553
QY	3001	CCCAAGGGCGAAAGTCATTTGGAATTTTAAATCATAGGGGAAAACAGTCTCTTCTAAA	3060
Db	3554	CCCAAGGGCGAAAGTCATTTGGAATTTTAAATCATAGGGGAAAACAGTCTCTTCTAAA	3613
QY	3061	TGCTCTTATCTTTTGGTTTGTCAAAAGAGAACTAAGAGCAGGACAGAGGCAACGT	3120
Db	3614	TGCTCTTATCTTTTGGTTTGTCAAAAGAGAACTAAGAGCAGGACAGAGGCAACGT	3673
QY	3121	GGAGGCTGAAAAACAGTGCAGAGAGCTTTGCAATGAGTCAGTAGCACAAAAGAGATGA	3180
Db	3674	GGAGGCTGAAAAACAGTGCAGAGAGCTTTGCAATGAGTCAGTAGCACAAAAGAGATGA	3733
QY	3181	CATTACTAGCACTATAAACCTGTGCTCTGAGAAAATGCTTCAATGATATAT	3240
Db	3734	CATTACTAGCACTATAAACCTGTGCTCTGAGAAAATGCTTCAATGATATAT	3793
QY	3241	GTGACTATTTTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTTAATAGAAATCCCAA	3300
Db	3794	GTGACTATTTTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTTAATAGAAATCCCAA	3853
QY	3301	TTTTCAGGAGTGGTGTGTCAATAAACCGCTCTGTGGCCAGTGTAAAGAAAAA	3353
Db	3854	TTTTCAGGAGTGGTGTGTCAATAAACCGCTCTGTGGCCAGTGTAAAGAAAAA	3906
RESULT 11			
AX470095			
LOCUS			

RESULT 11
AX470095
LOCUS

..CATGTGCTGCCCTCAGCGTCTCTCCATCTCACCGGCAAGTACGTCCCAACCAAC 991
J1 ACCTACACCAACATGAGAACTGCTCTCGCCCTCTGCGAGGCAACAGCAGAGAGCGG 420
992 ACCTACACCAACATGAGAACTGCTCTCGCCCTCTGCGAGGCAACAGCAGAGAGCGG 1051
421 ACCTTTGCGGTGACCTCAATAGCACTGCTACCGGACAGCTTTCTTCCGGAATATCTT 480
1052 ACCTTTGCGGTGACCTCAATAGCACTGCTACCGGACAGCTTTCTTCCGGAATATCTT 1111
481 AATGATACACACGGCTCTTACGTGCAACCGGCTGGAAGAGTGGGTGCGACTCTTTAA 540
1112 AATGATACACACGGCTCTTACGTGCAACCGGCTGGAAGAGTGGGTGCGACTCTTTAA 1171
541 AACTCCGCTTTTATAACTACACGCTGTGTGGAAGCGGGTGAAGAAAGACACGCTCC 600
1172 AACTCCGCTTTTATAACTACACGCTGTGTGGAAGCGGGTGAAGAAAGACACGCTCC 1231
601 GACTACTCCAAAGATTAACCTACAGACCTCATACCAATGACAGCTGAGCTTTCCGC 660
1232 GACTACTCCAAAGATTAACCTACAGACCTCATACCAATGACAGCTGAGCTTTCCGC 1291
661 ACCTCCAAAGATTAACCTACAGACCTCATACCAATGACAGCTGAGCTTTCCGC 720
1292 ACCTCCAAAGATTAACCTACAGACCTCATACCAATGACAGCTGAGCTTTCCGC 1351
721 CACGCGCTGAGGATTAACCTACAGACCTCATACCAATGACAGCTGAGCTTTCCGC 780
1352 CACGCGCTGAGGATTAACCTACAGACCTCATACCAATGACAGCTGAGCTTTCCGC 1411
781 ATCACGCGGAGTACAACTACGCGGCAACCGGAGCAAACTGATGATGCGGTACAG 840
1412 ATCACGCGGAGTACAACTACGCGGCAACCGGAGCAAACTGATGATGCGGTACAG 1471
841 GGGCCCATGAGCCCATTCACATGGAATTCACCAATGCTTCCAGCGGAAGCGTTGCA 900
1472 GGGCCCATGAGCCCATTCACATGGAATTCACCAATGCTTCCAGCGGAAGCGTTGCA 1531
901 ACCCTCATGTGCGTGAAGCTCCATGAGAGAGATTTACACATGCTGTTGAGAGCGG 960
1532 ACCCTCATGTGCGTGAAGCTCCATGAGAGAGATTTACACATGCTGTTGAGAGCGG 1591
961 GAGCTGGAACAACGATACATGATATACACCGGACCAAGCTTACACATCGGCCAGTTT 1020
1592 GAGCTGGAACAACGATACATGATATACACCGGACCAAGCTTACACATCGGCCAGTTT 1651
1021 GGCCTGTGAAAGGGAATTCATGCAATGATGATTTGACATCAGGTTCCGTTCTACGT 1080
1652 GGCCTGTGAAAGGGAATTCATGCAATGATGATTTGACATCAGGTTCCGTTCTACGT 1711
1081 AGGGGCCCCAAGTGAAGCGGCTGCTGATCCCAATGCTGCTCAACATGACCTG 1140
1712 AGGGGCCCCAAGTGAAGCGGCTGCTGATCCCAATGCTGCTCAACATGACCTG 1771
1141 GCGCCCAACCATCTTGACATTTGAGGCTGGAATACCTGCGGATATGGAACGGAAATCC 1200
1772 GCGCCCAACCATCTTGACATTTGAGGCTGGAATACCTGCGGATATGGAACGGAAATCC 1831
1201 ATCTCAAGCTGCTGACACGAGCGGCTGATCGGTTTCACTTGAAGAGAGATG 1260
1832 ATCTCAAGCTGCTGACACGAGCGGCTGATCGGTTTCACTTGAAGAGAGATG 1891
1261 AGGGTCTGGGGAGCTCTTCTTGTGTGAGAGGCAAGCTGCTACACAGAGACAAT 1320
1892 AGGGTCTGGGGAGCTCTTCTTGTGTGAGAGGCAAGCTGCTACACAGAGACAAT 1951
1321 GACAAGTGGACGCGCCAGGAGGAACTTTCTGCGCAAGTACAGCGTGTGAAGACCTG 1380
1952 GACAAGTGGACGCGCCAGGAGGAACTTTCTGCGCAAGTACAGCGTGTGAAGACCTG 2011
1381 TGTACAGCTGCTGATACACAGCGGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 1440
2012 TGTACAGCTGCTGATACACAGCGGCTGTGAGCAGCTGGGACAGAGTGGCAGTGTGTG 2071

1441 GAGGACCGCACGCGGAAGCTGAAGCTGATAGTGAAGGGCCCATGCGGTGTGGCGGC 1500
2072 GAGGACCGCACGCGGAAGCTGAAGCTGATAGTGAAGGGCCCATGCGGTGTGGCGGC 2131
1501 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTACGGGAGGGGACGAGGCGCTCAC 1560
2132 AGCAGAGCCCTCTCCAACTCTGTCGCAAGTACTACGGGAGGGGACGAGGCGCTCAC 2191
1561 TGTGACAGCGGGAATCAAGCTCAGCTCGCGGAGCGCGGAAAAAACTCTTCAAGAG 1620
2192 TGTGACAGCGGGAATCAAGCTCAGCTCGCGGAGCGCGGAAAAAACTCTTCAAGAG 2251
1621 AAGTACAGCGGCGAGTATGTCGCGAGTCTCATCGCTCAGTGGCCATCGAGGTGAG 1680
2252 AAGTACAGCGGCGAGTATGTCGCGAGTCTCATCGCTCAGTGGCCATCGAGGTGAG 2311
1681 GGCAGGGGTGACAGGTAGGCTTGGGTGATGCGCGCCAGCGCCGAAAACTCTCAAGCGG 1740
2312 GGCAGGGGTGACAGGTAGGCTTGGGTGATGCGCGCCAGCGCCGAAAACTCTCAAGCGG 2371
1741 CACTGGCCAGCGGCGCTTGAAGACCAAGATGACAGGATGGTGGGACTTTCAAGTGGCACT 1800
2372 CACTGGCCAGCGGCGCTTGAAGACCAAGATGACAGGATGGTGGGACTTTCAAGTGGCACT 2431
1801 GAGGCGCTTCCGAGTACTCAGCGCGCAACCCCAATTAAGTGAACATCGGTGCTACATC 1860
2432 GAGGCGCTTCCGAGTACTCAGCGCGCAACCCCAATTAAGTGAACATCGGTGCTACATC 2491
1861 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAACTCCCTGAGGCGCTT 1920
2492 CTAGAGAACGACACAGTCCAGTGTGACCTGGAACCTGTACAACTCCCTGAGGCGCTT 2551
1921 GACCAAGAGTGCACATCGACACCAAGATTAAGAACTCTGAGAAACAAATTAAGAACCTG 1980
2552 GACCAAGAGTGCACATCGACACCAAGATTAAGAACTCTGAGAAACAAATTAAGAACCTG 2611
1981 AGGAGCTCCGAGCTCAGCTGAGAAAGCGCCAGAGATGCTGCTCAAAATC 2040
2612 AGGAGCTCCGAGCTCAGCTGAGAAAGCGCCAGAGATGCTGCTCAAAATC 2671
2041 AGCTACACACACCGGCGCTCAAGCGGCTCAAGCAAGAGGCTCCAGTCTGATCTTTC 2100
2672 AGCTACACACACCGGCGCTCAAGCGGCTCAAGCAAGAGGCTCCAGTCTGATCTTTC 2731
2101 AGGAGGCGCTGCAAGAGAGAGGACAGGCTGCTGCGGAGCAGAGGCGCAAGAG 2160
2732 AGGAGGCGCTGCAAGAGAGAGGACAGGCTGCTGCGGAGCAGAGGCGCAAGAG 2791
2161 AAATCCGCAAGCTGCTCAAGCGGCTGAGAAACAGACACGTCGAGCATGCGGCGCTC 2220
2792 AAATCCGCAAGCTGCTCAAGCGGCTGAGAAACAGACACGTCGAGCATGCGGCGCTC 2851
2221 AGGTGCTTCAACCGACACACCGAGCTGCGGAGCGGCGCTTCTGAGACACTGGGCGCT 2280
2852 AGGTGCTTCAACCGACACACCGAGCTGCGGAGCGGCGCTTCTGAGACACTGGGCGCT 2911
2281 TTTCTGCTGCTGACAGCGGCGCAACATTAAGCTGCTGCTGAGCATGAGCATCAATGAG 2340
2912 TTTCTGCTGCTGACAGCGGCGCAACATTAAGCTGCTGCTGAGCATGAGCATCAATGAG 2971
2341 ACTCAAAATTTCTTCTGCTGCTGCTGCTGCTTCTTCTGAGTACTTTGATCTCAAC 2400
2972 ACTCAAAATTTCTTCTGCTGCTGCTGCTTCTTCTGAGTACTTTGATCTCAAC 3031
2401 ACAGACCCCTTACAGCTGATGAATGAGTGAACACACTGAGCAGGAGTGTCTCAACAG 2460
3032 ACAGACCCCTTACAGCTGATGAATGAGTGAACACACTGAGCAGGAGTGTCTCAACAG 3091
2461 CTACAGTACAGCTCATGAGAGCTGAGAGCTGAGAGTGAAGGGTTAAGAGTGTAAACCCCGG 2520
3092 CTACAGTACAGCTCATGAGAGCTGAGAGCTGAGAGTGAAGGGTTAAGAGTGTAAACCCCGG 3151

526	GTTCGACCTCTTTAAAACTCCCGCTTTTATATCAACGCTGTGTGGAACGGGGTGAAA	585
484	GTTCGACCTCTTTAAAACTCCCGCTTTTATAACTACACGCTGTGTGGAACGGGGTGAAA	543
586	GAATAAGCAGCGCTCCGACTACTCCAAGGATTAACCTCAAGACCTCATCAACAATGACAGC	645
544	GAATAAGCAGCGCTCCGACTACTCCAAGGATTAACCTCAAGACCTCATCAACAATGACAGC	603
646	GTGAGCTTCTTCCGACAGTCCCAAGAAGATGTACCAGCAGAGGCGAGTCCCTCATGTGTCATC	705
604	GTGAGCTTCTTCCGACAGTCCCAAGAAGATGTACCAGCAGAGGCGAGTCCCTCATGTGTCATC	663
706	AGCATGACGCCCCCAGCGCCCTTGAGGNTTACGCCCAACAATATTACAGCCCTTTCCTCCA	765
664	AGCATGACGCCCCCAGCGCCCTTGAGGNTTACGCCCAACAATATTACAGCCCTTTCCTCCA	723
766	AACGCAATCTAGCACATCAACGCCGAGCTACAATACGCGGCCCAACCGGACAAACTGG	825
724	AACGCAATCTAGCACATCAACGCCGAGCTACAATACGCGGCCCAACCGGACAAACTGG	783
826	ATCATGCGCTACACGGGGCCCATGAAGCCCATCCACATGGAATTTCAACAACTGCTCCAG	885
784	ATCATGCGCTACACGGGGCCCATGAAGCCCATCCACATGGAATTTCAACAACTGCTCCAG	843
886	CGGAAGCGCTTGCAGACCCCTCATGTCGGTGGACGACTTCATGAGAGCAGTTTAAACAATG	945
844	CGGAAGCGCTTGCAGACCCCTCATGTCGGTGGACGACTTCATGAGAGCAGTTTAAACAATG	903
946	CTGGTTGAGACGGCGGAGCTTGGACAAACAAGTATACGATATACACGCGGACACAGGTTAC	1005
904	CTGGTTGAGACGGCGGAGCTTGGACAAACAAGTATACGATATACACGCGGACACAGGTTAC	963
006	CACATCGGCCAGTTTGGCTGTGGTGAAGGAAATCCATGTCATATGAGTTTGAATCAGCAG	1065
964	CACATCGGCCAGTTTGGCTGTGGTGAAGGAAATCCATGTCATATGAGTTTGAATCAGCAG	1023
066	GTCCCGTTCTACGTGAGGGGCCCAACGTGGAAAGCGGCTGTCTGAATCCCCACATCGTC	1125
024	GTCCCGTTCTACGTGAGGGGCCCAACGTGGAAAGCGGCTGTCTGAATCCCCACATCGTC	1083
126	CTCAACATTGACCTGGGCCCCCACCATCTCTGGACATTTGACGGCTTGACATCTCGCGGAT	1185
084	CTCAACATTGACCTGGGCCCCCACCATCTCTGGACATTTGACGGCTTGACATCTCGCGGAT	1143
186	ATGAGCGGGAATCCATCTCTCAAGCTGTGACACGAGGCGCGGTGGAATCGGTTTCAC	1245
144	ATGAGCGGGAATCCATCTCTCAAGCTGTGACACGAGGCGCGGTGGAATCGGTTTCAC	1203
246	TTGAAAAAGAGATGAGGGGTCTGCGGGAATCTCTTCTTGTGTGAGAGAGCAAGCTGCTTA	1305
204	TTGAAAAAGAGATGAGGGGTCTGCGGGAATCTCTTCTTGTGTGAGAGAGCAAGCTGCTTA	1263
306	CACAAGAGAGACAATGACAAGGTGGACGCCCGCAGGAGGAGAACTTTCTGCCCCAAGTACACAG	1365
264	CACAAGAGAGACAATGACAAGGTGGACGCCCGCAGGAGGAGAACTTTCTGCCCCAAGTACACAG	1323
366	CGTGTGAAGGACCTGTGTGACGCTGTGAGTATACACAGCGCGTGTGACAGCTTGGGACAG	1425
324	CGTGTGAAGGACCTGTGTGACGCTGTGAGTATACACAGCGCGTGTGACAGCTTGGGACAG	1383
426	AAGTGGCAGTGTGTGAGGAGCGCAACGGGGAAGCTGAAGCTGATAGTGTGCAAGGCGCCC	1485
384	AAGTGGCAGTGTGTGAGGAGCGCAACGGGGAAGCTGAAGCTGATAGTGTGCAAGGCGCCC	1443
486	ATGCGGCTGGCGCAGCAGAGCCCTCTCCAACTCTGTGCCCCAAGTACTACGGGACAGGCG	1545
444	ATGCGGCTGGCGCAGCAGAGCCCTCTCCAACTCTGTGCCCCAAGTACTACGGGACAGGCG	1503
546	AGCAGGCGCTGCACCTGTGACAGCGGGAATCAAGCTCAGCCTGCGCCGAGACCGCGGAAA	1605
504	AGCAGGCGCTGCACCTGTGACAGCGGGAATCAAGCTCAGCCTGCGCCGAGACCGCGGAAA	1563
606	AAACTCTTCAAGAGAGATGAAGGCCAGCTATGTTCGACAGTTCGCTCCATCCGCTCAGTG	1665

<hr/>						
AX470097	AX470097	Sequence 3 from Patent WO02052019.	2610 bp	DNA	linear	PAT 09-AUG-2002
LOCUS	DEFINITION					
ACCESSION	Accession					
VERSION	Version					
SOURCE	Source					
ORGANISM	Organism					
REFERENCE	Reference					
AUTHORS	Authors					
TITLE	Title					
JOURNAL	Journal					
FEATURES	Features					
source						
BASE COUNT						
ORIGIN						
Query Match 77.5%; Score 2605.2; DB 6; Length 2610;						
Best Local Similarity 99.9%; Pseed No. 0;						
Matches 2607; Conservative 0; Mismatches 3; Indels 0; Gaps 0;						
QY	46	ATGGGCCCCCGAGCTGTGTCGTGTCGCCAATGTTTCCGCACTGTGTTCTCCCTGCTGGGT	105			
DB	1	ATGGGCCCCCGAGCTGTGTCGTGTCGCCAATGTTTCCGCACTGTGTTCTCCCTGCTGGGT	60			
QY	106	GGAAGCTCGGCCTTCTGTGCGACCACCGCTGAAAGCGAGGTTTGAGAGGACCGCAGG	165			
DB	61	GGAAGCTCGGCCTTCTGTGCGACCACCGCTGAAAGCGAGGTTTGAGAGGACCGCAGG	120			
QY	166	AACATCGGCCCCCAACTCATCTCTGTGTGTCGCGAACACAGGATGTGGAGCTGGGTTCC	225			
DB	121	AACATCGGCCCCCAACTCATCTCTGTGTGTCGCGAACACAGGATGTGGAGCTGGGTTCC	180			
QY	226	ATGCAGGTGATGAACAAGACCCCGCGCATCATGGAGCAGGCGCGGCGCACTTCATCAAC	285			
DB	181	ATGCAGGTGATGAACAAGACCCCGCGCATCATGGAGCAGGCGCGGCGCACTTCATCAAC	240			
QY	286	GCCTTGTGACACACCCATGTGTCGCCCTCACGCTCTTCCAATCTCTCACTGCGCAAGTAC	345			
DB	241	GCCTTGTGACACACCCATGTGTCGCCCTCACGCTCTTCCAATCTCTCACTGCGCAAGTAC	300			
QY	346	GTCCACACCAACCAACTTAGACCAACCAATGAGAATCTCTCTCGCCCTCTCTGGCAGGCA	405			
DB	301	GTCCACACCAACCAACTTAGACCAACCAATGAGAATCTCTCTCGCCCTCTCTGGCAGGCA	360			
QY	406	CAGCAGAGAGCGCACCTTTGCGGTGTACTCAATAGCACTGGTACCGGACAGCTTTC	465			
DB	361	CAGCAGAGAGCGCACCTTTGCGGTGTACTCAATAGCACTGGTACCGGACAGCTTTC	420			
QY	466	TTGGGAAGTATCTTAATGAATACAAACGGCTCTTACGTGCGACACCGGCTGGAAGGAGTGG	525			
DB	421	TTGGGAAGTATCTTAATGAATACAAACGGCTCTTACGTGCGACACCGGCTGGAAGGAGTGG	480			
QY	526	GTGCGACTCTCTTAAAACCTCCCGCTTTTATAACTACACGCTGTGTCGGACACGGGTGAA	585			
DB	481	GTGCGACTCTCTTAAAACCTCCCGCTTTTATAACTACACGCTGTGTCGGACACGGGTGAA	540			
QY	586	GAAAGCAACGGCTCGCATCTTCCAGGNTTACCTCTACAGACTCATACCAATGACAGC	645			
DB	541	GAAAGCAACGGCTCGCATCTTCCAGGNTTACCTCTACAGACTCATACCAATGACAGC	600			
QY	646	GTGAGCTCTTTCGCGACGTCGCAAGAGATGTAACCGCAGACGGCGAGTCCTCATGPGCATC	705			
DB	601	GTGAGCTCTTTCGCGACGTCGCAAGAGATGTAACCGCAGACGGCGAGTCCTCATGPGCATC	660			
QY	706	AGCATGACAGCCCCCAACCGGCCCTGAGGATTACGCCCCCAATAATTACAGGCTCTTCCCA	765			

Db 661 AGCCATGACGCCCCCAACGCGCCTGAGGATTACGCCCAACAATATTACGCGCTCTTCCCA 720
2y 766 AAGCATCTCAGACATACACGCGAGCTCAACTACGCGCCCAACCGGACAAACACTGG 825
Db 721 AAGCATCTCAGACATACACGCGAGCTCAACTACGCGCCCAACCGGACAAACACTGG 780
2y 826 ATCATCGCTTACACGCGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAG 885
Db 781 ATCATCGCTTACACGCGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAG 840
2y 886 CGAAGCGCTTGAGACCTCTATGTCGGTGGAGCTCATGTGAGACGATTTACAACTG 945
Db 841 CGAAGCGCTTGAGACCTCTATGTCGGTGGAGCTCATGTGAGACGATTTACAACTG 900
2y 946 CTGGTTGAGACGCGCGAGTGGACACATCATGTAACACGCGCCGACCGAGCTTAC 1005
Db 901 CTGGTTGAGACGCGCGAGTGGACACATCATGTAACACGCGCCGACCGAGCTTAC 960
2y 1006 CACATCGGCCAGTTTGGCCTGTGTAAGGGAATCCATGCCATATGATTTGACATCAGG 1065
Db 961 CACATCGGCCAGTTTGGCCTGTGTAAGGGAATCCATGCCATATGATTTGACATCAGG 1020
2y 1066 GTCCGCTTACCTGAGGGCCCCACGCTGGAGCGGCTGTCTGAATCCGCCACATCGTC 1125
Db 1021 GTCCGCTTACCTGAGGGCCCCACGCTGGAGCGGCTGTCTGAATCCGCCACATCGTC 1080
2y 1126 CTCAACATTCACCTGCGCCCAACATCTGGAATTCAGGCGCTGACATACCTCGGAT 1185
Db 1081 CTCAACATTCACCTGCGCCCAACATCTGGAATTCAGGCGCTGACATACCTCGGAT 1140
2y 1186 ATGGAACGGAATTCATCTCAAGCTGTGGACACGAGCGCGGTGATCGGTTTAC 1245
Db 1141 ATGGAACGGAATTCATCTCAAGCTGTGGACACGAGCGCGGTGATCGGTTTAC 1200
2y 1246 TTGAAAAGAGATGAGGCTGCGGGAATCTCTTGTGGAGAGGCAAGCTGCTA 1305
Db 1201 TTGAAAAGAGATGAGGCTGCGGGAATCTCTTGTGGAGAGGCAAGCTGCTA 1260
2y 1306 CACAAGAGACAAATGACAGGTGGACGCGCCAGAGAGAACTTTCTGCCAAGTACCAG 1365
Db 1261 CACAAGAGACAAATGACAGGTGGACGCGCCAGAGAGAACTTTCTGCCAAGTACCAG 1320
2y 1366 CGTGTGAGGACCTGTCTCAGCTGTCTGAGTACAGACGCGGTGTGAGCAGCTGGACAG 1425
Db 1321 CGTGTGAGGACCTGTCTCAGCTGTCTGAGTACAGACGCGGTGTGAGCAGCTGGACAG 1380
2y 1426 AAGTGGCAGTGTGTGGAGGACGCGCACGCGGGAAGCTGAAGCTGCATAAGTGCAGAGGCC 1485
Db 1381 AAGTGGCAGTGTGTGGAGGACGCGCACGCGGGAAGCTGAAGCTGCATAAGTGCAGAGGCC 1440
2y 1486 ATGGGCTGGGCGGACGAGCCCTCTCCACTCTGTCGCCCAAGTACTACGGCGAGGGC 1545
Db 1441 ATGGGCTGGGCGGACGAGCCCTCTCCACTCTGTCGCCCAAGTACTACGGCGAGGGC 1500
2y 1546 AGCAGGCGCTCAGCTGTGACAGCGGGGACTTACAGCTCAGCTGGCGGACCGCCGAAA 1605
Db 1501 AGCAGGCGCTCAGCTGTGACAGCGGGGACTTACAGCTCAGCTGGCGGACCGCCGAAA 1560
2y 1606 AAATCTTCAAGAGAGATCAAGGCGAGCTATGTCCGAGTGTCTCCATCCGCTCAGTG 1665
Db 1561 AAATCTTCAAGAGAGATCAAGGCGAGCTATGTCCGAGTGTCTCCATCCGCTCAGTG 1620
2y 1666 GCCATCGAGGTGGACGCGGAGGCTGTACAGCTAGGCTGTGATGCGGCCAGGCCCGCA 1725
Db 1621 GCCATCGAGGTGGACGCGGAGGCTGTACAGCTAGGCTGTGATGCGGCCAGGCCCGCA 1680
2y 1726 AACCTCACAGCGGCACTGCGCAGGGGCCCTGAGACCAAGATGCAAGATGATGTGG 1785
Db 1681 AACCTCACAGCGGCACTGCGCAGGGGCCCTGAGACCAAGATGCAAGATGATGTGG 1740
2y 1786 GACTTCAGTGGACCTGAGGCTTCCGACTACTCAGCGCCACCCATTAAGTGACA 1845
Db 1741 GACTTCAGTGGACCTGAGGCTTCCGACTACTCAGCGCCACCCATTAAGTGACA 1800

Qy 1846 CATCGTGTACATCTTAGAAGACGACACAGTCCAGTGTGACCTTGACCTGTGTACAAGTCC 1905
Db 1801 CATCGTGTACATCTTAGAAGACGACACAGTCCAGTGTGACCTGTGTACAAGTCC 1860
Qy 1906 CTGAGCGCTTGGAAAGACCAACAGTGTGACATGACCAACAGATTTGAAACCTTGCAGAAC 1965
Db 1861 CTGAGCGCTTGGAAAGACCAACAGTGTGACATGACCAACAGATTTGAAACCTTGCAGAAC 1920
Qy 1966 AAATTAAGAACCTTGAGGAGTCCGAGGTCACTGAAGAAAAGCGGCCAGAGAATGT 2025
Db 1921 AAATTAAGAACCTTGAGGAGTCCGAGGTCACTGAAGAAAAGCGGCCAGAGAATGT 1980
Qy 2026 GACTGTCAAAAAATCAGTACCAACACCAACAGCAAAAAGCGGCCCTCAGACACAGAGTCC 2085
Db 1981 GACTGTCAAAAAATCAGTACCAACACCAACAGCAAAAAGCGGCCCTCAGACACAGAGTCC 2040
Qy 2086 AGTGTGATCTTTCAGGAGGCGCTGCAAGAGAGGACAAAGGTGTGGCTGTTCGCGGAG 2145
Db 2041 AGTGTGATCTTTCAGGAGGCGCTGCAAGAGAGGACAAAGGTGTGGCTGTTCGCGGAG 2100
Qy 2146 CAGAAGCGCAAGAAACTCCGCAAGCTCTCAAGCGCTGAGCAACACGACACAGTGC 2205
Db 2101 CAGAAGCGCAAGAAACTCCGCAAGCTCTCAAGCGCTGAGCAACACGACACAGTGC 2160
Qy 2206 AGCATGCGAGCTCAGTGTTCACCCACGACAAACAGCACTGGCAGACGGCGCTTTC 2265
Db 2161 AGCATGCGAGCTCAGTGTTCACCCACGACAAACAGCACTGGCAGACGGCGCTTTC 2220
Qy 2266 TGGACATCTGGGCGCTTCTGTGCTGCAACGAGCGCCAAATAACACGTACTGGTGCATG 2325
Db 2221 TGGACATCTGGGCGCTTCTGTGCTGCAACGAGCGCCAAATAACACGTACTGGTGCATG 2280
Qy 2326 AGGACCATCAATAGAGACTCAATTTCTCTTGTGAATTTGCAACTGCTTCTTAGAG 2385
Db 2281 AGGACCATCAATAGAGACTCAATTTCTCTTGTGAATTTGCAACTGCTTCTTAGAG 2340
Qy 2386 TACTTTGATCTCAACACAGACCCCTTACGCTGATGATGAGTGAACACACTGCAAGCAGG 2445
Db 2341 TACTTTGATCTCAACACAGACCCCTTACGCTGATGATGAGTGAACACACTGCAAGCAGG 2400
Qy 2446 GATGTCTTCAACACAGCTACAGTACAGTATGAGCTGAGGAGCTGCAAGGGTTACAAG 2505
Db 2401 GATGTCTTCAACACAGCTACAGTACAGTATGAGCTGAGGAGCTGCAAGGGTTACAAG 2460
Qy 2506 CAGTGTAAACCCCGGACTCGAACAATGCACTGAGCTTAAAGATGAGGAAGCTATGAG 2565
Db 2461 CAGTGTAAACCCCGGACTCGAACAATGCACTGAGCTTAAAGATGAGGAAGCTATGAG 2520
Qy 2566 CAATACAGGCGAGTTTCAGCGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCAAATCA 2625
Db 2521 CAATACAGGCGAGTTTCAGCGCTCGAAGTGGCCAGAAATGAAGAGACCTTCTTCAAATCA 2580
Qy 2626 CTGGACAACTGTGGAGGCTGGAAGGT 2655
Db 2581 CTGGACAACTGTGGAGGCTGGAAGGT 2610

RESULT 15

BD127910
LOCUS
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
BD127910
ACCESSION
BD127910.1 GI:23222855
VERSION
JP 2002017375-A/3341.
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2559)
Ota.T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

TITLE
Primer for synthesizing full-length cDNA and use thereof
JOURNAL
Patent: JP 2002017375-A 3341 22-JAN-2002;

JOURNAL

HELIX RESEARCH INSTITUTE

HELIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)

PN JP 2002017375-A/3341

PD 22-JAN-2002

07-JUL-2000 JP 2000253172

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO

PI
THTST,
ISHII,

PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
CHUNYU CHU NGUYEN

SHINICHI KOUJIMA,
PT TETSUJI OTSUKI HIROSHI KOGA

PC
FI
1E15001 015001, H15001 NOGA

0130N15/09 007414/A7 007415/19 0130

10.

PC : C12P21/02.C12O1/68//C12P21/08.G06F17/30.C12N15/00.C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

 Location/Qualifiers |

FT CDS (133).

Location/Qualifiers

source	1. .2559
--------	----------

/organism="Homo sapiens"

```
/mol_type="genomic DNA"
```

```
/db_xref="taxon:9606"
```

BASE COUNT: 732 a 635 c 666 g 52

ORIGIN

Q10

Query Match	Similarity	99.94%	Pred. No. 0	Matches 2296	Conservative 0	Mismatches 2	Indels 0	Gaps 0
1055	TTGACATCAGGGTCCCGTTCTACGTGAGGGGCCCCCAACGTGGAGCGCGTGTCTGAATC	1114						
2	TGGACATCAGGGTCCCGTTCTACGTGAGGGGCCCCCAACGTGGAGCGCGTGTCTGAATC	61						
1115	CCACATCGTCTCTCAACATTTGACCTGGCCGCCACCACTCTGGACATTCAGGAGCTGGACA	1174						
62	CCACATCGTCTCTCAACATTTGACCTGGCCGCCACCACTCTGGACATTCAGGAGCTGGACA	121						
1175	TACCTGGGATATGACGGGAAATCCATCTCAAGCTGTGGACACGAGCGCGCGTGA	1234						
122	TACCTGGGATATGACGGGAAATCCATCTCAAGCTGTGGACACGAGCGCGCGTGA	181						
1235	ATCGGTTTCACTTGAAAGAAGATGAGGGTCTGGCGGACTCTCTTCTGGTGGAGAGAG	1294						
182	ATCGGTTTCACTTGAAAGAAGATGAGGGTCTGGCGGACTCTCTTCTGGTGGAGAGAG	241						
1295	GCAGCTGCTACACAAGAGAGACAAATGCAAGGTGGACGCCACGAGGAGAACTTTCTGC	1354						
242	GCAGCTGCTACACAAGAGAGACAAATGCAAGGTGGACGCCACGAGGAGAACTTTCTGC	301						
1355	CCAAGTACACAGCGTGTGAAGGACCTGTGTCAAGCGTGTGAGTACCAAGCGCGGTGTGAGC	1414						
302	CCAAGTACACAGCGTGTGAAGGACCTGTGTCAAGCGTGTGAGTACCAAGCGCGGTGTGAGC	361						
1415	AGCTGGGACAGAGTGGCAGTGTGTGAGAGCGCCACGGGAGAGCTGAAGCTGGATAAGT	1474						
362	AGCTGGGACAGAGTGGCAGTGTGTGAGAGCGCCACGGGAGAGCTGAAGCTGGATAAGT	421						
1475	GCAGGGGCCCCATCGGGCTGGCGGCGACAGAGCCCTCTCCAACTCTCGTGCCCAAGTACT	1534						
422	GCAGGGGCCCCATCGGGCTGGCGGCGACAGAGCCCTCTCCAACTCTCGTGCCCAAGTACT	481						
1535	ACGGCAGGGCAGCGAGGCTTGCACCTGTGTGACACGGGGACTCAAGCTCAGGCTGGCCG	1594						
482	ACGGCAGGGCAGCGAGGCTTGCACCTGTGTGACACGGGGACTCAAGCTCAGGCTGGCCG	541						
1595	GACGCCGAAAAAACTCTTCAAGAGAGTAGTCAAGGCCAGCTATGTCCGAGTGGCTCCCA	1654						
542	GACGCCGAAAAAACTCTTCAAGAGAGTAGTCAAGGCCAGCTATGTCCGAGTGGCTCCCA	601						
1655	TCCGCTCAGTGCCCATCGAGGTGACGGCAGGGGTGTACCACTAGGCGCTGGGTGATCGCG	1714						

QY	2795	CAACATGACAGATTCTGGAGGATAACACGACGAGCAGAGATAA	2854
Db	1742	CAACATGACAGATTCTGGAGGATAACACGACGAGCAGAGATAA	1801
QY	2855	TTTGCCCTGCTTTGCTTTGGATTATACCTCAGAGCTGCACAAATG	2914
Db	1802	TTTGCCCTGCTTTGCTTTGGATTATACCTCAGAGCTGCACAAATG	1861
QY	2915	ATCAAAAGTCAACCACTAACCCCTCCCCAGAGCTCACAAGGAAAAC	2974
Db	1862	ATCAAAAGTCAACCACTAACCCCTCCCCAGAGCTCACAAGGAAAAC	1921
QY	2975	CGAGAGAGATTTCCTTGGAAATTTCTCCAGGGCGAAAGTCATTG	3034
Db	1922	CGAGAGAGATTTCCTTGGAAATTTCTCCAGGGCGAAAGTCATTG	1981
QY	3035	TAGGGAAAAGCAGTCTCTTCTTAATCCTCTTATTTCTTTGGTTT	3094
Db	1982	TAGGGAAAAGCAGTCTCTTCTTAATCCTCTTATTTCTTTGGTTT	2041
QY	3095	ACTAGAAGCAGGACGAGGCGAACGTGGAGAGGCTGAAAACAGT	3154
Db	2042	ACTAGAAGCAGGACGAGGCGAACGTGGAGAGGCTGAAAACAGT	2101
QY	3155	ATGAGTCAGTACGACAAAAGAGATGACATTTTACCTAGCACTAT	3214
Db	2102	ATGAGTCAGTACGACAAAAGAGATGACATTTTACCTAGCACTAT	2161
QY	3215	GAAGAACTGCCTTCATTGTATATATGTGACTATTTACATGTAA	3274
Db	2162	GAAGAACTGCCTTCATTGTATATATGTGACTATTTACATGTAA	2221
QY	3275	TAGGGGAACCTTAAGAAATCCCAATTTTCAGAGTGGTGTCAATA	3334
Db	2222	TAGGGGAACCTTAAGAAATCCCAATTTTCAGAGTGGTGTCAATA	2281
QY	3335	GGCCAGTGTAAAGAAAA	3352
Db	2282	GGCCAGTGTAAAGAAAA	2299

Search completed: February 14, 2004, 17:54:31
Job time : 11919 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2004, 22:58:11 ; Search time 76 Seconds
(without alignments)
2396.876 Million cell updates/sec

Title: US-10-025-966A-6

Perfect score: 4750

Sequence: 1 MGPSLVLCILSATVSLILG.....PEMKRPPSKSLGQLWEGWG 870

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched:

801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Description
1	4750	100.0	870 12 US-10-182-951-5
2	4750	100.0	870 12 US-10-025-966A-6
3	4750	100.0	870 12 US-10-025-966A-15
4	4750	100.0	870 12 US-10-265-071-6
5	4750	100.0	870 12 US-10-265-071-15
6	4750	100.0	870 12 US-09-833-245-650
7	4746	99.9	870 12 US-09-970-287-2
8	4746	99.9	870 12 US-09-833-245-649
9	4746	99.9	870 12 US-10-264-237-2754
10	4746	99.9	870 12 US-09-970-287-12
11	4723.5	99.4	867 9 US-09-989-723-84
12	4723.5	99.4	867 9 US-09-989-723-84
13	4723.5	99.4	867 9 US-09-989-727-84
14	4723.5	99.4	867 9 US-09-989-727-84
15	4723.5	99.4	867 10 US-09-989-731-84

16	4723.5	99.4	867 10 US-09-989-732-84	Sequence 84, Appl
17	4723.5	99.4	867 10 US-09-991-073-84	Sequence 84, Appl
18	4723.5	99.4	867 10 US-09-990-442-84	Sequence 84, Appl
19	4723.5	99.4	867 10 US-09-991-163-84	Sequence 84, Appl
20	4723.5	99.4	867 10 US-09-993-604-84	Sequence 84, Appl
21	4723.5	99.4	867 10 US-09-990-456-84	Sequence 84, Appl
22	4723.5	99.4	867 10 US-09-989-721-84	Sequence 84, Appl
23	4723.5	99.4	867 10 US-09-992-598-84	Sequence 84, Appl
24	4723.5	99.4	867 10 US-09-989-293A-84	Sequence 84, Appl
25	4723.5	99.4	867 10 US-09-989-735-84	Sequence 84, Appl
26	4723.5	99.4	867 10 US-09-990-444-84	Sequence 84, Appl
27	4723.5	99.4	867 10 US-09-991-181-84	Sequence 84, Appl
28	4723.5	99.4	867 10 US-09-989-730-84	Sequence 84, Appl
29	4723.5	99.4	867 10 US-09-990-436-84	Sequence 84, Appl
30	4723.5	99.4	867 10 US-09-993-687-84	Sequence 84, Appl
31	4723.5	99.4	867 11 US-09-989-734-84	Sequence 84, Appl
32	4723.5	99.4	867 11 US-09-997-653-84	Sequence 84, Appl
33	4723.5	99.4	867 11 US-09-993-667-84	Sequence 84, Appl
34	4723.5	99.4	867 11 US-09-997-428-84	Sequence 84, Appl
35	4723.5	99.4	867 11 US-09-997-666-84	Sequence 84, Appl
36	4723.5	99.4	867 11 US-09-990-438-84	Sequence 84, Appl
37	4723.5	99.4	867 11 US-09-990-562-84	Sequence 84, Appl
38	4723.5	99.4	867 11 US-09-990-711-84	Sequence 84, Appl
39	4723.5	99.4	867 11 US-09-989-726-84	Sequence 84, Appl
40	4723.5	99.4	867 11 US-09-998-156-84	Sequence 84, Appl
41	4723.5	99.4	867 11 US-09-990-437-84	Sequence 84, Appl
42	4723.5	99.4	867 11 US-09-991-157-84	Sequence 84, Appl
43	4723.5	99.4	867 11 US-09-997-514-84	Sequence 84, Appl
44	4723.5	99.4	867 11 US-09-997-573-84	Sequence 84, Appl
45	4723.5	99.4	867 11 US-09-991-172-84	Sequence 84, Appl

ALIGNMENTS

RESULT 1

US-10-182-951-5
; Sequence 5, Application US/10182951
; Publication NO. US20030138895A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Aameena R.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SHIH, Leo L.
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: FI-0033 PCT
; CURRENT APPLICATION NUMBER: US/10/182,951
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/181,856; 10/183,684; 60/185,141; 60/186,818; 60/188,345
; PRIOR FILING DATE: 2000-02-11; 2000-02-17; 2000-02-25; 2000-03-03; 2000-03-09; 2000-03-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030138895A1 1558210CDI
US-10-182-951-5

Query Match 100.0%; Score 4750; DB 12; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60
Db 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTGYVHNHNTYNNENCSSPSWQA 120
Db 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTGYVHNHNTYNNENCSSPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVVPGKKEWGLLKNSRFYNTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVVPGKKEWGLLKNSRFYNTLCRNGVK 180
QY 181 EKXGSDYSKDYLTDLTNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
Db 181 EKXGSDYSKDYLTDLTNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
QY 241 NASOHTTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHTTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNTIYVYTADHGHIQGFGLVKGKSMPEYFDIRVPPYVGRPNVEAGCLNPHIV 360
Db 301 LVETGELDNTIYVYTADHGHIQGFGLVKGKSMPEYFDIRVPPYVGRPNVEAGCLNPHIV 360
QY 361 LNIDLAFTILDIAGLDIPADMDGKSIKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
Db 361 LNIDLAFTILDIAGLDIPADMDGKSIKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
QY 421 HKRDNDKVDQAQENFLPKYQKVLDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAQENFLPKYQKVLDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV 540
QY 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
QY 601 HRCYIILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLRVGRHGLKKRPEEC 660
Db 601 HRCYIILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLRVGRHGLKKRPEEC 660
QY 661 DCHKISYHTQHKGLKXRGSSLHPPFKGLQEKDKVWLLREQKXKGLKRLKRLQNNDT 720
Db 661 DCHKISYHTQHKGLKXRGSSLHPPFKGLQEKDKVWLLREQKXKGLKRLKRLQNNDT 720
QY 721 SMPGLTCFTHDNQHWOTAPPTWTLGPFCACTSANNNTYWCWRTINETHNLFCEFATGFE 780
Db 721 SMPGLTCFTHDNQHWOTAPPTWTLGPFCACTSANNNTYWCWRTINETHNLFCEFATGFE 780
QY 781 YFDLNTDPYQLMANAVNTLDRDVLNQLHVQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMANAVNTLDRDVLNQLHVQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGQWEGWEG 870
```

RESULT 2

```
US-10-025-966A-6
; Sequence 6, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; APPLICANT: Stefan, Hemmerich
; TITLE OF INVENTION: Sulfatases and methods of use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025, 966A
```

```
; CURRENT FILING DATE: 2002-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-025-966A-6
```

```
Query Match 100.0%; Score 4750; DB 12; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60
Db 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTGYVHNHNTYNNENCSSPSWQA 120
Db 61 MQVMKTRIMEGGGAHFNAFVTTMCCPSRSSILTGYVHNHNTYNNENCSSPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVVPGKKEWGLLKNSRFYNTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVVPGKKEWGLLKNSRFYNTLCRNGVK 180
QY 181 EKXGSDYSKDYLTDLTNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
Db 181 EKXGSDYSKDYLTDLTNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
QY 241 NASOHTTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHTTPSYNAPNDKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNTIYVYTADHGHIQGFGLVKGKSMPEYFDIRVPPYVGRPNVEAGCLNPHIV 360
Db 301 LVETGELDNTIYVYTADHGHIQGFGLVKGKSMPEYFDIRVPPYVGRPNVEAGCLNPHIV 360
QY 361 LNIDLAFTILDIAGLDIPADMDGKSIKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
Db 361 LNIDLAFTILDIAGLDIPADMDGKSIKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
QY 421 HKRDNDKVDQAQENFLPKYQKVLDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAQENFLPKYQKVLDLQRAEYQTAQCEQLGQKQWQVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV 540
QY 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVHVGLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
QY 601 HRCYIILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLRVGRHGLKKRPEEC 660
Db 601 HRCYIILENTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLRVGRHGLKKRPEEC 660
QY 661 DCHKISYHTQHKGLKXRGSSLHPPFKGLQEKDKVWLLREQKXKGLKRLKRLQNNDT 720
Db 661 DCHKISYHTQHKGLKXRGSSLHPPFKGLQEKDKVWLLREQKXKGLKRLKRLQNNDT 720
QY 721 SMPGLTCFTHDNQHWOTAPPTWTLGPFCACTSANNNTYWCWRTINETHNLFCEFATGFE 780
Db 721 SMPGLTCFTHDNQHWOTAPPTWTLGPFCACTSANNNTYWCWRTINETHNLFCEFATGFE 780
QY 781 YFDLNTDPYQLMANAVNTLDRDVLNQLHVQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMANAVNTLDRDVLNQLHVQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
```

QY 841 QYRQORRKWPEMKRPPSSKSLGQLWEGWEG 870
DB 841 QYRQORRKWPEMKRPPSSKSLGQLWEGWEG 870

RESULT 3
US-10-025-966A-15
; Sequence 15, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; APPLICANT: Stefan, Hemmerich
; TITLE OF INVENTION: Sulfatases and methods of use
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025,966A
; PRIOR FILING DATE: 2002-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-966A-15

Query Match 100.0%; Score 4750; DB 12; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPSLVLCLSATVFSLLGGSSAFLSHRLKGFQDRNRNIRNIIILVLTDDQVELGS 60
DB 1 MGPSLVLCLSATVFSLLGGSSAFLSHRLKGFQDRNRNIRNIIILVLTDDQVELGS 60
QY 61 MQVWKTRIMEQGAHFINAFVTPMCCPSRSSILGKYVHNHTYNNENCSPPSWQA 120
DB 61 MQVWKTRIMEQGAHFINAFVTPMCCPSRSSILGKYVHNHTYNNENCSPPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRFFNYTLCRNGVK 180
DB 121 QHESRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGMVLSHAAPHGPEPSAPOYSRLFP 240
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGMVLSHAAPHGPEPSAPOYSRLFP 240
QY 241 NASQHITPSYNYAPNPKHWMRYTGPMPKHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASQHITPSYNYAPNPKHWMRYTGPMPKHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNTIIVYTDHGYHIGQFGLVKGSMPYEDIRVPPFVYRGNVEAGCLNPHIV 360
DB 301 LVETGELDNTIIVYTDHGYHIGQFGLVKGSMPYEDIRVPPFVYRGNVEAGCLNPHIV 360
QY 361 LNIIDLAPTILDIAGLIDIPADMDGSIILKLDTERPVNRFLKGMVLSHAAPHGPEPSAPOYSRLFP 420
DB 361 LNIIDLAPTILDIAGLIDIPADMDGSIILKLDTERPVNRFLKGMVLSHAAPHGPEPSAPOYSRLFP 420
QY 421 HKRNDKVDAGEENFLPKYQVVKDLCQRAEYQACEQLGQKWCQVEDATCKLKLHCKGK 480
DB 421 HKRNDKVDAGEENFLPKYQVVKDLCQRAEYQACEQLGQKWCQVEDATCKLKLHCKGK 480
QY 481 MRLGGSRLSRLNLPKYQVVKDLCQRAEYQACEQLGQKWCQVEDATCKLKLHCKGK 540
DB 481 MRLGGSRLSRLNLPKYQVVKDLCQRAEYQACEQLGQKWCQVEDATCKLKLHCKGK 540
QY 541 AIEVDGRVYHVLGDAQAQPNRLTKRHWFGAPEDQDDQDGGDFSGTGGLPDYSAAPIKVT 600
DB 541 AIEVDGRVYHVLGDAQAQPNRLTKRHWFGAPEDQDDQDGGDFSGTGGLPDYSAAPIKVT 600

QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETLQNKIKNLRVGRHLKKRPPSEC 660
DB 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETLQNKIKNLRVGRHLKKRPPSEC 660
QY 661 DCHKISYHTQHKGRGLKRGSSLHPPKGLQEKQVWLLRBOQRKQKGLKRLKRLQNDTC 720
DB 661 DCHKISYHTQHKGRGLKRGSSLHPPKGLQEKQVWLLRBOQRKQKGLKRLKRLQNDTC 720
QY 721 SMPGLTCTFHDNQHWQTAPFTWLTGPFCACTSANNTYWCMTINETHNLFCEFAIGLE 780
DB 721 SMPGLTCTFHDNQHWQTAPFTWLTGPFCACTSANNTYWCMTINETHNLFCEFAIGLE 780
QY 781 YFDLNTDPPQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPPQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRNMDLGLKDGGSYE 840
QY 841 QYRQORRKWPEMKRPPSSKSLGQLWEGWEG 870
DB 841 QYRQORRKWPEMKRPPSSKSLGQLWEGWEG 870

RESULT 4
US-10-265-071-6
; Sequence 6, Application US/10265071
; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-265-071-6

Query Match 100.0%; Score 4750; DB 12; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPSLVLCLSATVFSLLGGSSAFLSHRLKGFQDRNRNIRNIIILVLTDDQVELGS 60
DB 1 MGPSLVLCLSATVFSLLGGSSAFLSHRLKGFQDRNRNIRNIIILVLTDDQVELGS 60
QY 61 MQVWKTRIMEQGAHFINAFVTPMCCPSRSSILGKYVHNHTYNNENCSPPSWQA 120
DB 61 MQVWKTRIMEQGAHFINAFVTPMCCPSRSSILGKYVHNHTYNNENCSPPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRFFNYTLCRNGVK 180
DB 121 QHESRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGMVLSHAAPHGPEPSAPOYSRLFP 240
DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRLKGMVLSHAAPHGPEPSAPOYSRLFP 240
QY 241 NASQHITPSYNYAPNPKHWMRYTGPMPKHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASQHITPSYNYAPNPKHWMRYTGPMPKHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNTIIVYTDHGYHIGQFGLVKGSMPYEDIRVPPFVYRGNVEAGCLNPHIV 360

Db 301 LVETGELDNITYIYVYADHGVHIGQFGLVKGKSMPEYEFDIRVFPYVGRPNVZAGCLNPHIV 360
Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRWVWDSFLVERGKLL 420
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRWVWDSFLVERGKLL 420
Qy 421 HKRDNDKVAQAEENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVAQAEENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLKHKCKGP 480
Qy 481 MRLGGSRALSNLVPKYGGSEACTCDSDYKLSLAGREKKLFKKYKASVYRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYGGSEACTCDSDYKLSLAGREKKLFKKYKASVYRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYIILENDTVQCDLDLYKSLQAWKDKHLHIHDEITLQNKIKNIREVRGHLKKRPEEC 660
Db 601 HRCYIILENDTVQCDLDLYKSLQAWKDKHLHIHDEITLQNKIKNIREVRGHLKKRPEEC 660
Qy 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKKLAKLKLRLQNNDT 720
Db 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKKLAKLKLRLQNNDT 720
Qy 721 SMPGLTCFTHDNOHWOTAPFWTLGPFCACTSANNNTYWCRTINETHNPLFCFATGPLE 780
Db 721 SMPGLTCFTHDNOHWOTAPFWTLGPFCACTSANNNTYWCRTINETHNPLFCFATGPLE 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVOLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVOLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840

RESULT 5

US-10-265-071-15
; Sequence 15, Application US/10265071
; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hammerich, Stefan
; APPLICANT: Ionica, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-265-071-15

Query Match 100.0%; Score 4750; DB 12; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGPPSLVLCLSATVPSLLGGSSAFSLHRLKGRFQRDRNRIRPNIIIVLTDQDQVEIGS 60
Db 1 MGPPSLVLCLSATVPSLLGGSSAFSLHRLKGRFQRDRNRIRPNIIIVLTDQDQVEIGS 60

Qy 61 MQVANKTRRIMEGGAHFINAFVTTMPCPSRSSILTKYVHNHNTYTNNENCSSPSWOA 120
Db 61 MQVANKTRRIMEGGAHFINAFVTTMPCPSRSSILTKYVHNHNTYTNNENCSSPSWOA 120
Qy 121 QHSRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNSRFPNYTLCNGVK 180
Db 121 QHSRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNSRFPNYTLCNGVK 180
Qy 181 EKSGSDYSKDYLTDLITNDSVSFRTSKMYPHRPVLVWISHAAPHGPDSDAPQYSLRLEP 240
Db 181 EKSGSDYSKDYLTDLITNDSVSFRTSKMYPHRPVLVWISHAAPHGPDSDAPQYSLRLEP 240
Qy 241 NASCHITPSTNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
Db 241 NASCHITPSTNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELDNITYIYVYADHGVHIGQFGLVKGKSMPEYEFDIRVFPYVGRPNVZAGCLNPHIV 360
Db 301 LVETGELDNITYIYVYADHGVHIGQFGLVKGKSMPEYEFDIRVFPYVGRPNVZAGCLNPHIV 360
Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRWVWDSFLVERGKLL 420
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRWVWDSFLVERGKLL 420
Qy 421 HKRDNDKVAQAEENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVAQAEENFLPKYQVVKDLQORAEYQTAQCEQLGQKQVCVEDATGKLKHKCKGP 480
Qy 481 MRLGGSRALSNLVPKYGGSEACTCDSDYKLSLAGREKKLFKKYKASVYRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYGGSEACTCDSDYKLSLAGREKKLFKKYKASVYRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYIILENDTVQCDLDLYKSLQAWKDKHLHIHDEITLQNKIKNIREVRGHLKKRPEEC 660
Db 601 HRCYIILENDTVQCDLDLYKSLQAWKDKHLHIHDEITLQNKIKNIREVRGHLKKRPEEC 660
Qy 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKKLAKLKLRLQNNDT 720
Db 661 DCHKISYHTQHKGRKXHRGSSLHPFRKGLQEKDKVWLLREQKKLAKLKLRLQNNDT 720
Qy 721 SMPGLTCFTHDNOHWOTAPFWTLGPFCACTSANNNTYWCRTINETHNPLFCFATGPLE 780
Db 721 SMPGLTCFTHDNOHWOTAPFWTLGPFCACTSANNNTYWCRTINETHNPLFCFATGPLE 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVOLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVOLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870

RESULT 6

US-09-833-245-650
; Sequence 650, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384

; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 2267
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 650
 ; LENGTH: 870
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-833-245-650

Query Match 100.0%; Score 4750; DB 12; Length 870;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRRNRIRNIIILVLTDDQDVELGS 60
 DB 1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRRNRIRNIIILVLTDDQDVELGS 60
 QY 61 MQVMKTRRIMEQGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120
 DB 61 MQVMKTRRIMEQGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120
 QY 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
 DB 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
 QY 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHPGPDSSAPQYSLFP 240
 DB 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHPGPDSSAPQYSLFP 240
 QY 241 NASQHITPSYNAVPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
 DB 241 NASQHITPSYNAVPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
 QY 301 LVEGTGELDNITYIVYTADHGYHIGQGLVKGSMPEYFDIRVPFYVGPVNEAGCLNPHIV 360
 DB 301 LVEGTGELDNITYIVYTADHGYHIGQGLVKGSMPEYFDIRVPFYVGPVNEAGCLNPHIV 360
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420
 DB 361 LNIDLAPTILDIAGLDIPADMDGKSILKLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420
 QY 421 HKRDNKDVAQAEENFLPKYQVQDLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
 DB 421 HKRDNKDVAQAEENFLPKYQVQDLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVRSRSTRSV 540
 DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVRSRSTRSV 540
 QY 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
 DB 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
 QY 601 HRCYILENDTVQCDDLDLYKSLQAWKDKHLHDHIEITLQNKIKNLREVRGHLKKRPEEC 780
 DB 601 HRCYILENDTVQCDDLDLYKSLQAWKDKHLHDHIEITLQNKIKNLREVRGHLKKRPEEC 780
 QY 781 YFDLNTDPYQLMNAVNTLDRVLNQLHVLQVLMELRSCGKYQCNPRTRNMDLGLKDGSGYE 840
 DB 781 YFDLNTDPYQLMNAVNTLDRVLNQLHVLQVLMELRSCGKYQCNPRTRNMDLGLKDGSGYE 840
 QY 841 QYRQFQRKWPMEKRPSSKSLGQLWEGWEG 870
 DB 841 QYRQFQRKWPMEKRPSSKSLGQLWEGWEG 870

RESULT 7
 US-09-970-287-2
 ; Sequence 2, Application US/09970287
 ; Publication No. US20030166283A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GLUCKSMANN, Maria A.
 ; APPLICANT: MEYERS, Rachel
 ; APPLICANT: KAPELLER-LIBERMANN, Rosana
 ; APPLICANT: SILOS-SANTIAGO, Inmaculada
 ; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR
 ; FILE REFERENCE: 10147-6IU1
 ; CURRENT APPLICATION NUMBER: US/09/970,287
 ; CURRENT FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: US 60/257,082
 ; PRIOR FILING DATE: 2000-12-21
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 870
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-970-287-2

Query Match 99.9%; Score 4746; DB 12; Length 870;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRRNRIRNIIILVLTDDQDVELGS 60
 DB 1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRRNRIRNIIILVLTDDQDVELGS 60
 QY 61 MQVMKTRRIMEQGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120
 DB 61 MQVMKTRRIMEQGGAHFNAFVITPMCCPSRSSILTGKYVHNHTYNNENCSPPSWQA 120
 QY 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
 DB 121 QHESITFAVYLNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
 QY 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHPGPDSSAPQYSLFP 240
 DB 181 EKHGSDYKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHPGPDSSAPQYSLFP 240
 QY 241 NASQHITPSYNAVPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
 DB 241 NASQHITPSYNAVPNDKHWIMRYTGMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
 QY 301 LVEGTGELDNITYIVYTADHGYHIGQGLVKGSMPEYFDIRVPFYVGPVNEAGCLNPHIV 360
 DB 301 LVEGTGELDNITYIVYTADHGYHIGQGLVKGSMPEYFDIRVPFYVGPVNEAGCLNPHIV 360
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420
 DB 361 LNIDLAPTILDIAGLDIPADMDGKSILKLDTERPVNRPHLKKMRVWRDSFLVERGKLL 420
 QY 421 HKRDNKDVAQAEENFLPKYQVQDLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
 DB 421 HKRDNKDVAQAEENFLPKYQVQDLCORAEYQTAQCEQLGQKWCQVEDATGKLLHKCKGP 480
 QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVRSRSTRSV 540
 DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASVRSRSTRSV 540
 QY 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
 DB 541 AIEVDGRVYHVLGDAAPRNLTKRHPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
 QY 601 HRCYILENDTVQCDDLDLYKSLQAWKDKHLHDHIEITLQNKIKNLREVRGHLKKRPEEC 660
 DB 601 HRCYILENDTVQCDDLDLYKSLQAWKDKHLHDHIEITLQNKIKNLREVRGHLKKRPEEC 660
 QY 661 DCHKISYHTQHKGRKLRGSSSLHPFRKQLQEKDKVLLRBEQFRKKLRLKLLKRLQNDTC 720

Db 661 DCHKISYHTQHGRLKRGSSLPFRKGLQEKDKVLLRQKQKRLKRLQNNDTG 720
QY 721 SMPGLTCTFTHDQHWQTAPFWTLGPFCACTSANNTYKCMETINETHFLCFEATGFL 780
Db 721 SMPGLTCTFTHDQHWQTAPFWTLGPFCACTSANNTYKCMETINETHFLCFEATGFL 780
QY 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
RESULT 8
US-09-833-245-649
; Sequence 649, Application US/09833245
; Publication No. US200400101341
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 649
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (534)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-649
Query Match 99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60
Db 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60
QY 61 MQVMNKTREIMEQGAHFNAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSWQA 120
Db 61 MQVMNKTREIMEQGAHFNAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLNKSRFYNTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLNKSRFYNTLCRNGVK 180
QY 181 EKHGSDYSKDYLTDLTNDVSFFRTSKOMYPHRVLVMSHAAPHGSDSAPQYSLRFP 240
Db 181 EKHGSDYSKDYLTDLTNDVSFFRTSKOMYPHRVLVMSHAAPHGSDSAPQYSLRFP 240
QY 241 NASQHTPSYNTAPNDKHWIRYTGPMKPIHMEFTNMLQRKLOTLMSSVDDSMETIYNN 300
Db 241 NASQHTPSYNTAPNDKHWIRYTGPMKPIHMEFTNMLQRKLOTLMSSVDDSMETIYNN 300
QY 301 LVETGELDNTIYVYADHGYHIGQGLVKGKMPYEFDRVPFYRGPNVEAGCLNPHIV 360
Db 301 LVETGELDNTIYVYADHGYHIGQGLVKGKMPYEFDRVPFYRGPNVEAGCLNPHIV 360
QY 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVNRFLKQKVRWRDPSFLVERGKLL 420

Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVNRFLKQKVRWRDPSFLVERGKLL 420
QY 421 HKRDNDKVDQAQENFLPKYQVRKDIQRAEYQTACEQLGQKQWCVEDATGKLKHKCKGP 480
Db 421 HKRDNDKVDQAQENFLPKYQVRKDIQRAEYQTACEQLGQKQWCVEDATGKLKHKCKGP 480
QY 481 MRIGGSRALSNLVPKYQGGSEACTCDSDGYKSLAGRRKLLFKKKYKASYVRSRSIRSV 540
Db 481 MRIGGSRALSNLVPKYQGGSEACTCDSDGYKSLAGRRKLLFKKKYKASYVRSRSIRSV 540
QY 541 ALEVGRVYHVGIGDAAPRNLTKEHWPAGPDQDDKGGSPSGTGGPLDYSAANPIKVT 600
Db 541 ALEVGRVYHVGIGDAAPRNLTKEHWPAGPDQDDKGGSPSGTGGPLDYSAANPIKVT 600
QY 601 HRCYILENDTVQCCLDLKYSLOAKDKHLHIDHEIETLQNKIKNLRVGRHLKGRPSEC 660
Db 601 HRCYILENDTVQCCLDLKYSLOAKDKHLHIDHEIETLQNKIKNLRVGRHLKGRPSEC 660
QY 661 DCHKISYHTQHGRLKRGSSLPFRKGLQEKDKVLLRQKQKRLKRLQNNDTG 720
Db 661 DCHKISYHTQHGRLKRGSSLPFRKGLQEKDKVLLRQKQKRLKRLQNNDTG 720
QY 721 SMPGLTCTFTHDQHWQTAPFWTLGPFCACTSANNTYKCMETINETHFLCFEATGFL 780
Db 721 SMPGLTCTFTHDQHWQTAPFWTLGPFCACTSANNTYKCMETINETHFLCFEATGFL 780
QY 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
RESULT 9
US-10-264-237-2754
; Sequence 2754, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2754
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (534)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-2754
Query Match 99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60
Db 1 MGPPSLVLCCLLSATVPSLLGGSSAFSLSHRLKGRFQDRRNRIRPNILVLTDDQDVELGS 60
QY 61 MQVMNKTREIMEQGAHFNAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSWQA 120
Db 61 MQVMNKTREIMEQGAHFNAFVTTMCCPSSSILTKGYVHNHTYNNENCSPPSWQA 120

QY 121 QHESRTFAVILNSTGYRTAFKGYLNEYNGSYVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
DB 121 QHESRTFAVILNSTGYRTAFKGYLNEYNGSYVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
QY 181 EKHSYDYSKDYLTDLIINDSVSPFRTSKMYPHRPVLMVISHAAPHGPDSPQYSLFP 240
DB 181 EKHSYDYSKDYLTDLIINDSVSPFRTSKMYPHRPVLMVISHAAPHGPDSPQYSLFP 240
QY 241 NASQHIITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNN 300
DB 241 NASQHIITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNLYIVVTADHGVIHGOGLVKGKSNPYEFDIRVPPVYRGNVVEACLNPHIV 360
DB 301 LVETGELDNLYIVVTADHGVIHGOGLVKGKSNPYEFDIRVPPVYRGNVVEACLNPHIV 360
QY 361 LNIDLAFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKKRVWRDSFLVERGKLL 420
DB 361 LNIDLAFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKKRVWRDSFLVERGKLL 420
QY 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGQKQCVEDATGKLLHKCKGP 480
DB 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGQKQCVEDATGKLLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKLLFKKKYKASYVRSRISV 540
DB 481 MRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKLLFKKKYKASYVRSRISV 540
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
QY 601 HRCVILENDTVQCDLDLYKSLQAWKHLDHEIETLQNKIKNLREVGRHLKCKRPEEC 660
DB 601 HRCVILENDTVQCDLDLYKSLQAWKHLDHEIETLQNKIKNLREVGRHLKCKRPEEC 660
QY 661 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNDT 720
DB 661 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNDT 720
QY 721 SMPGLTCFTHDNQHWQAPFWTLGPPFCACTSANNNTYWCMTINETHNFCEFATGFLE 780
DB 721 SMPGLTCFTHDNQHWQAPFWTLGPPFCACTSANNNTYWCMTINETHNFCEFATGFLE 780
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFQRRKWPENKRPSSKSLGQWEGWEG 870
DB 841 QYRQFQRRKWPENKRPSSKSLGQWEGWEG 870

RESULT 10

US-09-970-287-12
; Sequence 12, Application US/09970287
; Publication No. US2003016283A1
; GENERAL INFORMATION:
; APPLICANT: MEYERS, Rachel
; APPLICANT: KAPLELLER-LIBERMAN, Rosana
; APPLICANT: SILOS-SANTIAGO, Inmaculada
; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR
; FILE REFERENCE: 10147-61U1
; CURRENT APPLICATION NUMBER: US/09/970,287
; PRIOR FILING DATE: 2001-10-03
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 885

TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-287-12

Query Match 99.9%; Score 4746; DB 12; Length 885;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGPSLVLCILSATVFSLLGGSSAFLSHRLKGRFQDRRNRIRNIIILVLTDDQDVELGS 60
DB 16 MGPSLVLCILSATVFSLLGGSSAFLSHRLKGRFQDRRNRIRNIIILVLTDDQDVELGS 75
QY 61 MOVNKKTRIRIEOGGAHFINAVFTPMCCPSRSSRILTGKYVHNENTYNNENCSPPQWA 120
DB 76 MOVNKKTRIRIEOGGTHFINAVFTPMCCPSRSSRILTGKYVHNENTYNNENCSPPQWA 135
QY 121 QHESRTFAVILNSTGYRTAFKGYLNEYNGSYVPPGKWEVGLLKNRFRFYNTLCRNGVK 180
DB 136 QHESRTFAVILNSTGYRTAFKGYLNEYNGSYVPPGKWEVGLLKNRFRFYNTLCRNGVK 195
QY 181 EKHSYDYSKDYLTDLIINDSVSPFRTSKMYPHRPVLMVISHAAPHGPDSPQYSLFP 240
DB 196 EKHSYDYSKDYLTDLIINDSVSPFRTSKMYPHRPVLMVISHAAPHGPDSPQYSLFP 255
QY 241 NASQHIITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNN 300
DB 256 NASQHIITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNN 315
QY 301 LVETGELDNLYIVVTADHGVIHGOGLVKGKSNPYEFDIRVPPVYRGNVVEACLNPHIV 360
DB 316 LVETGELDNLYIVVTADHGVIHGOGLVKGKSNPYEFDIRVPPVYRGNVVEACLNPHIV 375
QY 361 LNIDLAFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKKRVWRDSFLVERGKLL 420
DB 376 LNIDLAFTILDIAGLDIPADMDGKSIILKLDTERPPVNRPHLKKKRVWRDSFLVERGKLL 435
QY 421 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGQKQCVEDATGKLLHKCKGP 480
DB 436 HKRDNKVDQAEENFLPKYQVVKDLQCORAEYQVACEQLGQKQCVEDATGKLLHKCKGP 495
QY 481 MRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKLLFKKKYKASYVRSRISV 540
DB 496 MRLGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKLLFKKKYKASYVRSRISV 555
QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 556 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 615
QY 601 HRCVILENDTVQCDLDLYKSLQAWKHLDHEIETLQNKIKNLREVGRHLKCKRPEEC 660
DB 616 HRCVILENDTVQCDLDLYKSLQAWKHLDHEIETLQNKIKNLREVGRHLKCKRPEEC 675
QY 661 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNDT 720
DB 676 DCHKISYHTQHKGLKRGSSLHPFRKGLQEKDKVLLREQKQKLLKLLKRLQNNDT 735
QY 721 SMPGLTCFTHDNQHWQAPFWTLGPPFCACTSANNNTYWCMTINETHNFCEFATGFLE 780
DB 736 SMPGLTCFTHDNQHWQAPFWTLGPPFCACTSANNNTYWCMTINETHNFCEFATGFLE 795
QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
DB 796 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 855
QY 841 QYRQFQRRKWPENKRPSSKSLGQWEGWEG 870
DB 856 QYRQFQRRKWPENKRPSSKSLGQWEGWEG 885

RESULT 11

US-09-989-722-84
; Sequence 84, Application US/09989722
; Patent No. US20020072067A1

GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C63
CURRENT APPLICATION NUMBER: US/09/989,722
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089601
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 99.48; Score 4723.5; DB 9; Length 867;
Best Local Similarity 99.78; Pred. No. 0;
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60
DB 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRDRNIRPNILVLTDDQDVELGS 60
QY 61 MOVNKKTRIMEOGGAHFINAFVTPMCCPSRSSILTKYVHNHTYTNNCSSPSWQA 120
DB 61 MOVNKKTRIMEOGGAHFINAFVTPMCCPSRSSILTKYVHNHTYTNNCSSPSWQA 120
QY 121 QHESRTFAVLNSTGYRTAFPGKLYNEINGSYVPPGKWKVGLLKNSRFNYTLCEGVK 180
DB 121 QHESRTFAVLNSTGYRTAFPGKLYNEINGSYVPPGKWKVGLLKNSRFNYTLCEGVK 180
QY 181 EKGSDYSKDYLTDLITNDSVSFFRSKQYPRHPVLMVISHAHPGSDSAPQYSRLFP 240
DB 181 EKGSDYSKDYLTDLITNDSVSFFRSKQYPRHPVLMVISHAHPGSDSAPQYSRLFP 240

QY 241 NASQHITPSYNYAPNDKHWIMRYTGPMPKPIHMEFTNNMQSKRLQTLMSVDDSMETIYNN 300
DB 241 NASQHITPSYNYAPNDKHWIMRYTGPMPKPIHMEFTNNMQSKRLQTLMSVDDSMETIYNN 300
QY 301 LVETGELDNTYIVVTADHGXYHIGQGLVKGKSMPEYEDIRVPFYVRGENVEAGCINPHIV 360
DB 301 LVETGELDNTYIVVTADHGXYHIGQGLVKGKSMPEYEDIRVPFYVRGENVEAGCINPHIV 360
QY 361 LNIDLAPTILDIAGLDIPADMDGKSIKLKLDTERPVRNPHLKKKQVRWRDSFLVERGKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSIKLKLDTERPVRNPHLKKKQVRWRDSFLVERGKLL 420
QY 421 HKRNDKVDQAEENFLPKYQKVLDLCORAEYOTACEOLGQKQWQVEDATGKLHKCKGP 480
DB 421 HKRNDKVDQAEENFLPKYQKVLDLCORAEYOTACEOLGQKQWQVEDATGKLHKCKGP 480
QY 481 MRLGGSRALSNLVPKYTQGGSEACTCDSGDYKLSLAGRRKKLKKKQKASVYRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYTQGGSEACTCDSGDYKLSLAGRRKKLKKKQKASVYRSRSIRSV 540
QY 541 AIEVDGRVYHVLGDAAPRNLTGRHWPAGPEDDODKGGDFSGTGGGLPDYSAANPIKVT 600
DB 541 AIEVDGRVYHVLGDAAPRNLTGRHWPAGPEDDODKGGDFSGTGGGLPDYSAANPIKVT 600
QY 601 HRCVILENDTVQCDLDLYKSLQAWKDKHLHDHIEITLQNKIKNLRVVRGHLKKRPEEC 660
DB 601 HRCVILENDTVQCDLDLYKSLQAWKDKHLHDHIEITLQNKIKNLRVVRGHLKKRPEEC 660
QY 661 DCHKISYHTQHKRLKHGSSLHPFRKGLQEKDKVLLRQKRRKKLKLKRLQNNNDTC 720
DB 661 DCHKISYHTQHKRLKHGSSLHPFRKGLQEKDKVLLRQKRRKKLKLKRLQNNNDTC 720
QY 721 SMPGLTCFTHDNQHWQAPFWTLGPPCACSANNNTYWCMTTNETHNEFCBEATGFLE 780
DB 721 SMPGLTCFTHDNQHWQAPFWTLGPPCACSANNNTYWCMTTNETHNEFCBEATGFLE 780
QY 781 YFDLNTDPYQLMNAVNTLDRDLNLQHLVQLMELRSCKYKQCNPRTNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYQLMNAVNTLDRDLNLQHLVQLMELRSCKYKQCNPRTNMDLGLKDGGSYE 840
QY 841 QYRQFQRKWPMPKRPSSKSLGQLWEGWEG 870
DB 838 QYRQFQRKWPMPKRPSSKSLGQLWEGWEG 867

RESULT 12
US-09-989-723-84
; Sequence 84, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaudo, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1062
CURRENT APPLICATION NUMBER: US/09/989,723
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 99.4%; Score 4723.5; DB 9; Length 867;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

2y 1 MGPPSLVLCILSATVPSLILGGSFAFSLHRLKGRFQDRDRNIRPNILVLTDDQVELGS 60
2b 1 MGPPSLVLCILSATVPSLILGGSFAFSLHRLKGRFQDRDRNIRPNILVLTDDQVELGS 60

2y 61 MQVWNTKRIIMEOGGAHFINAFVTTMPCPCSSSILTKYVHNHNTYNNENCSFSPQQA 120
2b 61 MQVWNTKRIIMEOGGAHFINAFVTTMPCPCSSSILTKYVHNHNTYNNENCSFSPQQA 120

2y 121 QHSRTFAVNLSTGYRTAFPGKYLNEYNGSVYPPGKWKWGLLKNRPYNTLCRNGVK 180
2b 121 QHSRTFAVNLSTGYRTAFPGKYLNEYNGSVYPPGKWKWGLLKNRPYNTLCRNGVK 180

2y 181 EKHGSDYKDYLDLTNDVSFRTSKMYRHPVLMVISHAAPHGPDSDAPQYSRLFP 240
2b 181 EKHGSDYKDYLDLTNDVSFRTSKMYRHPVLMVISHAAPHGPDSDAPQYSRLFP 240

2y 241 NASCHITPSYNAFNPDKHWMRYTGPMPKIHMFTNMLQRKQLTMSVDDSMETIYNN 300
2b 241 NASCHITPSYNAFNPDKHWMRYTGPMPKIHMFTNMLQRKQLTMSVDDSMETIYNN 300

2y 301 LVETGELDNTIYVTADHGHVIGFGLVKGKSNPYEFDIRVPPYVRGNVAGCLNPHIV 360
2b 301 LVETGELDNTIYVTADHGHVIGFGLVKGKSNPYEFDIRVPPYVRGNVAGCLNPHIV 360

2y 361 LNIDLAPTILDIAGLIDIPADMDGKSLIKLIDTFRPVNRFHLKKQKVRVWDSFLVERGKLL 420
2b 361 LNIDLAPTILDIAGLIDIPADMDGKSLIKLIDTFRPVNRFHLKKQKVRVWDSFLVERGKLL 420

2y 421 HKRDNKVDQAEENFLPKYQVRVXDLQRAEYQACBOLGQKQCVEDATGKLLHKCKGP 480
2b 421 HKRDNKVDQAEENFLPKYQVRVXDLQRAEYQACBOLGQKQCVEDATGKLLHKCKGP 480

2y 481 MRLGGSRALSNLVPKYQGSEACTCDSDGYKLSLAGRKKLFKKYKASYVRSRSIRSV 540
2b 481 MRLGGSRALSNLVPKYQGSEACTCDSDGYKLSLAGRKKLFKKYKASYVRSRSIRSV 540

2y 541 AIEVDGRVHVGLGDAAPQNLTKRWHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
2b 541 AIEVDGRVHVGLGDAAPQNLTKRWHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

Qy 601 HRCVILENDTVQCDLDBLYKSLQAWKHKLHIDHEIETLQNKIKNLRVFGHLKKKPEEC 660
Db 601 HRCVILENDTVQCDLDBLYKSLQAWKHKLHIDHEIETLQNKIKNLRVFGHLKKKPEEC 660

Qy 661 DCHKISYHTQHKGRGLXHRGSSLHPFRKGLQEKDKVLLREQKKKLRKLLKELQNDTC 720
Db 661 DCHKISYHTQHKGRGLXHRGSSLHPFRKGLQEKDKVLLREQKKKLRKLLKELQNDTC 720

Qy 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSAANNNTYWCMTINETHNLFCEFATGFLE 780
Db 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSAANNNTYWCMTINETHNLFCEFATGFLE 780

Qy 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVOLWEURSCGYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDYPQLMNAVNTLDRDVLNQLHVOLWEURSCGYKQCNPRTRNMDLGLKDGGSYE 837

Qy 841 QYRQFQRRKWPENKMPSPSSKSLGQWEGWEG 870
Db 838 QYRQFQRRKWPENKMPSPSSKSLGQWEGWEG 867

RESULT 13

US-09-989-279-84
; Sequence 84, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Garber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07

PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089801
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089908
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089948
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089952
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090246
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090252
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090254
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090349
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090355
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090429
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090431
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090435
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090444
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090445
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090472
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090557
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090676
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090678
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090690
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090694
PRIOR FILING DATE:	1998-06-26
PRIOR APPLICATION NUMBER:	60/090695
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090696
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090862
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091544
PRIOR FILING DATE:	1998-07-01
PRIOR APPLICATION NUMBER:	60/091519
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091626
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091633
PRIOR FILING DATE:	1998-07-02
PRIOR APPLICATION NUMBER:	60/091978
PRIOR FILING DATE:	1998-07-0-

PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match	99.4%	Score 4723.5	DB 9	Length 867
Best Local Similarity	99.7%	Pred. No. 0	Mismatches 0	Indels 3
Matches 867	Conservative	0		Gaps 1
1	MGPPSLVLCCLLSATVFLSGLGSSAFLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS	60		
1	MGPPSLVLCCLLSATVFLSGLGSSAFLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS	60		
61	MOVNKRTRIMEGGGAHFNATVTPMCCPSRSSILTCKYVHNHTYNNENCSPSQA	120		
61	MOVNKRTRIMEGGGAHFNATVTPMCCPSRSSILTCKYVHNHTYNNENCSPSQA	120		
121	QHSRTFAVYLNSTGYRTAFFGKYLNEYNGSVPPGKWEVGLLKNRSFYNYTKCRNGVK	180		
121	QHSRTFAVYLNSTGYRTAFFGKYLNEYNGSVPPGKWEVGLLKNRSFYNYTKCRNGVK	180		
181	EXHGSYSDXYTDLITNDSVFFRTSKWYHPHVPVLVISHAAPHGPEDSAPQYSRLFP	240		
181	EXHGSYSDXYTDLITNDSVFFRTSKWYHPHVPVLVISHAAPHGPEDSAPQYSRLFP	240		
241	NASOHITPSYNYAPNPDKHWIRYTGPKPIHMEFTNMLQRKLOTLSVDDSMETIYNN	300		
241	NASOHITPSYNYAPNPDKHWIRYTGPKPIHMEFTNMLQRKLOTLSVDDSMETIYNN	300		
301	LVTGELDNITYVYTADHGYH:GQGLVKGKSMPEYDIPVPFYVRGPNVAGCLNPHIV	360		
301	LVTGELDNITYVYTADHGYH:GQGLVKGKSMPEYDIPVPFYVRGPNVAGCLNPHIV	360		
361	LNIDLAPTITLDIAGLDIPADMDGKSTILKLDTERPVNRPFLKXKRWVDSFLVERGKLL	420		
361	LNIDLAPTITLDIAGLDIPADMDGKSTILKLDTERPVNRPFLKXKRWVDSFLVERGKLL	420		
421	HKRDNKDVAQENFLPKYQVKDLCORAEYQACEQLGQKQWQCVEDATGKLLHKCKGP	480		
421	HKRDNKDVAQENFLPKYQVKDLCORAEYQACEQLGQKQWQCVEDATGKLLHKCKGP	480		
481	MRLGSRALSNLVPKYGGSSACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV	540		
481	MRLGSRALSNLVPKYGGSSACTCDSDGYKLSLAGRRKKLFKKYKASVYRSRSIRSV	540		
541	AIEVDGRVTHVGLDAAQPRNLTQKHPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT	600		
541	AIEVDGRVTHVGLDAAQPRNLTQKHPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT	600		
601	HRCYILENDTVQCDLDLYKSLQAWKDLHIDHIEITLQNKIKNLREVRGHLKKGKPEEC	660		
601	HRCYILENDTVQCDLDLYKSLQAWKDLHIDHIEITLQNKIKNLREVRGHLKKGKPEEC	660		
661	DCHKISYHTQKGLKGRGSSLHPKGLQKDKWLLREOKRKKLKLKRLQNNDTIC	720		
661	DCHKISYHTQKGLKGRGSSLHPKGLQKDKWLLREOKRKKLKLKRLQNNDTIC	720		
721	SNPGLTCFTHDNQHWQAPFWTLGPFCACTSANNNTYWCWRTINETHNFLFCEFATGFL	780		
721	SNPGLTCFTHDNQHWQAPFWTLGPFCACTSANNNTYWCWRTINETHNFLFCEFATGFL	780		
781	YFDLNTDPTQLMNAVNTLDRDVLNQLHVQMLRSCKYKQCNPRTRMNDGLKDGGSYE	840		
781	YFDLNTDPTQLMNAVNTLDRDVLNQLHVQMLRSCKYKQCNPRTRMNDL---DGSYE	837		
841	QYRQFORRWKPMKRPSSKSLGQWEGWG	870		
838	QYRQFORRWKPMKRPSSKSLGQWEGWG	867		

RESULT 14
US-09-989-727-84
; Sequence 84, Application US/09989727

Patent No. US20020072497A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC65
CURRENT APPLICATION NUMBER: US/09/989,727
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088859
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09
;;
Query Match 99.4%; Score 4723.5; DB 9; Length 867;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MGPPSLVLCLLSATVFSLLGGSSAFLSHRLKGRFQDRRNIRNIIILVLTDDQDELGS 60
Db 1 MGPPSLVLCLLSATVFSLLGGSSAFLSHRLKGRFQDRRNIRNIIILVLTDDQDELGS 60

QY 61 MQVMNKTTRIMEQGAHFNAFVTPMCCPSRSSILTKGYVHNHTYNNENCSPPSWQA 120
Db 61 MQVMNKTTRIMEQGAHFNAFVTPMCCPSRSSILTKGYVHNHTYNNENCSPPSWQA 120

QY 121 QHESRTFAVLNLTGRTAFKGYLNEVNGSVPPGKWKVGLLKNSRFYNTILCRNGVK 180
Db 121 QHESRTFAVLNLTGRTAFKGYLNEVNGSVPPGKWKVGLLKNSRFYNTILCRNGVK 180

QY 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPYSRLFP 240

241 NASQHTPSYNTAPNDKWIWRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
241 NASQHTPSYNTAPNDKWIWRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
301 LVETGELDNIVYVYADHGVHIGQFGLVKGKSMPEYFEDIRVFPYVVRGPNVAGCLNPHIV 360
301 LVETGELDNIVYVYADHGVHIGQFGLVKGKSMPEYFEDIRVFPYVVRGPNVAGCLNPHIV 360
361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXKXVVRWDSFLVERGKLL 420
361 LNIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKXKXVVRWDSFLVERGKLL 420
421 HKEDNDKVDQAENFUPKYQVVKDLCORAEYQTAECQLGQKWCQVEDATGKLKHKCKGP 480
421 HKEDNDKVDQAENFUPKYQVVKDLCORAEYQTAECQLGQKWCQVEDATGKLKHKCKGP 480
481 MRLGGSRALSNLVPKYGGSEACTCSDGYKLSLAGREKLLFKKXKASVYRSRSISRV 540
481 MRLGGSRALSNLVPKYGGSEACTCSDGYKLSLAGREKLLFKKXKASVYRSRSISRV 540
541 AIEVDGRVYHVGGLDAAQPRNLTKRHWPGAPEDODDKGGDFSGTGLPDYSAANPIKVT 600
541 AIEVDGRVYHVGGLDAAQPRNLTKRHWPGAPEDODDKGGDFSGTGLPDYSAANPIKVT 600
601 HRCYILENDTVQCDLKYSLQAKDKHLHIDHIEITLQNKILNREVGRHLKXKRPBEC 660
601 HRCYILENDTVQCDLKYSLQAKDKHLHIDHIEITLQNKILNREVGRHLKXKRPBEC 660
661 DCHKISYHTQHKRLKRGSSLHPPFRKGLQKDKVWLLREQKXKXKLLKRLKRLQNNDT 720
661 DCHKISYHTQHKRLKRGSSLHPPFRKGLQKDKVWLLREQKXKXKLLKRLKRLQNNDT 720
721 SWPGLTCFTHDNOHWQTAFFWTLPFCACCTSANNTYWCMTINETHNPLFCBPATGILE 780
721 SWPGLTCFTHDNOHWQTAFFWTLPFCACCTSANNTYWCMTINETHNPLFCBPATGILE 780
781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQMLRSCGYKQCNPRTRNMDLGLKDGGSYE 840
781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQMLRSCGYKQCNPRTRNMDLGLKDGGSYE 840
841 OYRQFORRWKPEKRPSSKSLQGLWEGWEG 870
838 OYRQFORRWKPEKRPSSKSLQGLWEGWEG 867

RESULT 15

US-09-989-731-84

Sequence 84, Application US/09989731

Patent No. US20020103125A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas P.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same

FILE REFERENCE: P2730P1C70

CURRENT APPLICATION NUMBER: US/09/989,731

CURRENT FILING DATE: 2001-11-20

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088028

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088202

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088217

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088655

PRIOR FILING DATE: 1998-06-09

PRIOR APPLICATION NUMBER: 60/088734

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088738

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088742

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10

;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690

;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 99.4%; Score 4723.5; DB 10; Length 867;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 867; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MGPPSLVLCLLSATVFLGGSSAFLSHRLKGRFDRNRINRPNILVLTDQDVELGS 60
Db 1 MGPPSLVLCLLSATVFLGGSSAFLSHRLKGRFDRNRINRPNILVLTDQDVELGS 60

QY 61 MOVNKTTRRINEOGGAHFINAVTTMCCPSRSILTKGVVHNHTYNNENCSPPSWQA 120
Db 61 MOVNKTTRRINEOGGAHFINAVTTMCCPSRSILTKGVVHNHTYNNENCSPPSWQA 120

QY 121 QHESRTFAVLNSTGYRTAFPGKYLNEYNGSYVPPGKWEVGLLKNRSFYNYTLCRNGVK 180
Db 121 QHESRTFAVLNSTGYRTAFPGKYLNEYNGSYVPPGKWEVGLLKNRSFYNYTLCRNGVK 180

QY 181 EKHGSDYSKDYLTDLITNDSVSFRTSKQWPHRPVLMVISHAAPHGPDSPAQYSELFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFRTSKQWPHRPVLMVISHAAPHGPDSPAQYSELFP 240

QY 241 NASQHITPSYNAAPNDKHWIMRYTGPMKPIHMEFTNMLORKLOTLMSVDDSMETIYNN 300
Db 241 NASQHITPSYNAAPNDKHWIMRYTGPMKPIHMEFTNMLORKLOTLMSVDDSMETIYNN 300

QY 301 LVETGELDNITYVVTADHGVIHQFGLVKGKSMYPEDIRVPFVYRGPNVEAGCLNPHIV 360
Db 301 LVETGELDNITYVVTADHGVIHQFGLVKGKSMYPEDIRVPFVYRGPNVEAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPYNRPHLKKMRVWRDSFLVERKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMDGKSIILKLDTERPYNRPHLKKMRVWRDSFLVERKLL 420

QY 421 HKRDNKVDQAEEENFLPKYQKVDLCORAEYQACEQLGQKWCQVEDATGKLHLKCKGP 480
Db 421 HKRDNKVDQAEEENFLPKYQKVDLCORAEYQACEQLGQKWCQVEDATGKLHLKCKGP 480

QY 481 MRLGGSRALNLPVKYKYGQSEACTCDSDYKLSLAGRRKLFKKYKASVYRSRSRVS 540
Db 481 MRLGGSRALNLPVKYKYGQSEACTCDSDYKLSLAGRRKLFKKYKASVYRSRSRVS 540

QY 541 AIEVDGRVYHVGLGDAAPRNLTKRHWPGAPEDQDDXGDDGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVGLGDAAPRNLTKRHWPGAPEDQDDXGDDGDFSGTGGLPDYSAANPIKVT 600

2Y 601 HRCYILENDTVQCDLDLYKSLQAWKDLHIDHIEITLQNKIKNLRVGRHLKKRPEEC 660
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 601 HRCYILENDTVQCDLDLYKSLQAWKDLHIDHIEITLQNKIKNLRVGRHLKKRPEEC 660
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 661 DCHKISYHTQHKGRKLRGSSSLHPRFKGLQEKDKVWLLRBOQRKKKLRKLRLQNNDT 720
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 661 DCHKISYHTQHKGRKLRGSSSLHPRFKGLQEKDKVWLLRBOQRKKKLRKLRLQNNDT 720
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 721 SMPGLTCFTHDNOHWOTAPFTWLGPPFACTSANNNTYWCMTINETHNLFCEFATGFE 780
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 721 SMPGLTCFTHDNOHWOTAPFTWLGPPFACTSANNNTYWCMTINETHNLFCEFATGFE 780
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 841 QYRQFORRKWPEMKRPSSKSLGQWEGWEG 870
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2Y 838 QYRQFORRKWPEMKRPSSKSLGQWEGWEG 867
2b ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: February 14, 2004, 23:07:49
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

Run on: February 14, 2004, 19:42:15 ; Search time 39 seconds
(without alignments)
1049.057 Million cell updates/sec

Title: US-10-025-966A-6
Effect score: 4750
Sequence: 1 MGPPSLVLCLSATVFSLLG.....PEMKRPSKSLGQWEGWG 870

Coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	858.5	18.1	552	1	GL6S HUMAN
2	856	18.0	559	1	GL6S_CAPHI
3	452.5	9.5	649	1	ARS VOLCA
4	405	8.5	647	1	ARS CHLE
5	302.5	6.4	497	1	YIDJ_ECOLI
6	260	5.5	512	1	BETC_RHIME
7	251	5.3	560	1	YDNN_ECOLI
8	238.5	5.0	583	1	STS HUMAN
9	238	5.0	567	1	ARS STRPU
10	234.5	4.9	550	1	IDS HUMAN
11	232	4.9	551	1	ARS HEMPU
12	226	4.8	563	1	IDS_MOUSE
13	218.5	4.6	522	1	GAGS_HUMAN
14	210.5	4.4	506	1	ARS_MOUSE
15	210.5	4.4	533	1	ARS_HUMAN
16	210.5	4.4	577	1	STS RAT
17	203	4.3	535	1	ARS FELCA
18	201.5	4.2	464	1	ARS_KLEAE
19	200	4.2	593	1	ARSD_HUMAN
20	198.5	4.2	591	1	ARSF_HUMAN
21	194.5	4.1	507	1	ARS_HUMAN
22	189.5	4.0	624	1	STS_MOUSE
23	183.5	3.9	589	1	ARS_HUMAN
24	176.5	3.7	502	1	SPHM_HUMAN
25	176	3.7	551	1	ASLA_ECOLI
26	170	3.6	473	1	ARS_MOUSE
27	155.5	3.3	535	1	ARS_PSEAE
28	138.5	2.9	510	1	GNPI_LISIN
29	137.5	2.9	510	1	GNPI_LISMO
30	135.5	2.9	510	1	GNPI_ANTSP
31	134	2.8	533	1	GNPI_ANASP
32	128	2.7	285	1	ARSB_MOUSE
33	127	2.7	532	1	GNPI_SINY3

RESULT 1
GL6S_HUMAN STANDARD; PRT; 552 AA.
AC P15586;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE N-acetylglucosamine-6-sulfatase precursor (EC 3.1.6.14) (G6S)
DE (Glucosamine-6-sulfatase).
GN GNS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Endothelial cells;
RX MEDLINE=93098807; PubMed=1463457;
RA Robertson D.A., Freeman C., Morris C.P., Hopwood J.J.;
RT "A cDNA clone for human glucosamine-6-sulfatase reveals differences between arylsulphatases and non-arylsulphatases.";
RL Biochem. J. 288:539-544 (1992).
[2]
RP SEQUENCE OF 178-552 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89061714; PubMed=3196333;
RA Robertson D.A., Freeman C., Nelson P.V., Morris C.P., Hopwood J.J.;
RT "Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase.";
RL Biochem. Biophys. Res. Commun. 157:218-224 (1988).
CC -|- CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate group of the N-acetyl-D-glucosamine 6-sulfate units of heparan sulfate and keratan sulfate.
CC -|- SUBCELLULAR LOCATION: Lysosomal.
CC -|- DISEASE: DEFECTS IN GNS RESULT IN THE ACCUMULATION OF PARTIALLY DEGRADED HEPARAN SULFATE IN LYSOSOMES CAUSING ORGANELLE, CELL AND TISSUE DISTORTION, ULTIMATELY LEADING TO THE LYSOSOMAL STORAGE DISORDER, MUCOPOLYSACCHARIDOSIS TYPE IIID (MPS-IIID) (ALSO KNOWN AS SANFILIPPO D SYNDROME).
CC -|- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z12173; CAA78164.1; --
CC PIR; S27164; KKHUGU.
CC Genew; HGNC:4422; GNS.
CC MIM; 252940; --
CC GO; GO:0008449; F:N-acetylglucosamine-6-sulfatase activity; TAS.
CC GO; GO:0006027; P:glycosaminoglycan catabolism; TAS.
CC InterPro; IPR000917; Sulfatase.
CC Fram; PF00884; Sulfatase; 1.
CC -----

ALIGNMENTS

DR PROSITE; PS00523; SULFATASE 1; 1.
DR PROSITE; PS00149; SULFATASE 2; 1.
GW Hydrolase; Glycoprotein; Lysosome; Signal; Mucopolysaccharidosis.
FT SIGNAL 1 36
FT CHAIN 37 552
FT MOD_RES 91 91
FT N-ACETYLGLUCOSAMINE-6-SULFATASE.
FT 2-AMINO-3-OXOPROPIONIC ACID (BY
FT SIMILARITY).
FT CARBOHYD 111 111
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 117 117
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 210 210
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 279 279
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 362 362
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 422 422
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 449 449
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 480 480
FT N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 552 AA; 62082 MW; 85094043F6E64468 CRC64;
Query Match 18.1%; Score 858.5; DB 1; Length 552;
Best Local Similarity 27.2%; Pred. No. 1.3e-54;
Matches 226; Conservative 76; Mismatches 174; Indels 355; Gaps 18;
4 PSVLCLLSAT--VFLSGSSAPLSHRLKGRFDRNRNIRPNIILVLTDDQDVELGSM 61
24 PALLLVGLGCLGVFGVAAGTR-----RNVVLLTDDQDVELGSM 64
62 QVNMKTRINEOGAHPINAFVTPPCSSRSILTKGVVHNTVYN--NENCSPSQW 119
65 TPLKTKALIGEMGTFFSSAYVSALCCPSRSLTKGPHNHVHVNNTLEGCSKSNQ 124
120 AQHESRTFAVLNS-TGYRTAFPGKLYNEING-----SYVPGKWEVGLLKNRFRNY 172
125 KIQEPNTPFALLASMGQYQFFAGKLYNEVGPADAGLEHVLPGWSYVALEKSNKYNY 184
173 TLRNGVKRKGSDYKDYLTIDNDSVFFRTSKMYPHRVLMVISHAAPHGPDSEA 232
185 TLENGKARKHGENYSVDYLTDLVANSVLDYKSNF---EPFFMIATPAPHSPTAA 241
233 PQYSLRFPNASQHTPTSYNAPNPDHWMRYT-GPMKPIHMEFTNNLQKRLQTLMSVD 291
242 PQYKARQNVAPRNKNFN-IHGTNKKWLRQAKTPTNWSIGFLDNARKRWQTLISVD 300
292 DSMETIYNMLVETGELDNTYIVYADHGHIQGLVKGSMYPYRDIRVFFYVRGNVE 351
301 DLVEKLVKRLFTGELNNTYIFTSDNGYHTGQFSLPIDKRLQYEFDIKVPVLLVVRGFIK 360
352 AGCLNPHVLNIDAPTLDIAGLDI-PADMDGKSLKLDLTERPVRNRRHLKKKRVVRD 410
361 PNQTSKMLVANIDGPTLIDAGVLDNKTQMDGMSLLPIL---RGASNL-----TWRS 410
411 SFLVERGKLLHKRDNKVDYDAEENFLPKYQVRKDLQORAEYQACEQLGQKQCVEDATG 470
411 DVLVE----- 415
471 KKLHKCKGPNRLGGSRALSNLVPKYVQGSSEACTCDSDYKLSLAGRRKCLKKKYKAS 530
416 -----YQGE----- 420
531 YVRSRSIRSAIEVDGRVYHVLGDAQAQPNRLTKRHWFGAPEDQDDKDGDFSGTGGLPD 590
421 -----RNV----- 424
591 YSAANPIKVTHRCVILENDTVQCDLDLYKSLQAWKHLDHIEFTLQNKILNREVRG 650
425 ----- 424
651 HLKKRPEECCHKI SYHTQHKGRLLKHRGSSLHPFRKGLQKDKVLLRQKGRKKLKL 710
425 -----DFTCPSL----- 422

711 LKRLQNNDCSMGLT-CPTHNDQHWOTAPFTLGPFCACTSANNNTYWCMTINETHNF 769
433 -----PVSQCP-----PDCVEDAYNTYACVRTMSALNWL 464
770 LFCEP-ATGFLFYFDLNTDPYQLMNAVNTLDRDLVNLQHLVQLMEIRSKG 818
465 QYCEFDQEVFEVYVNLTPDQITNIAKTIDPELLGKNYRLMLLQSCSG 515
RESULT 2
ID GL6S_CAPHI STANDARD; PRT; 559 AA.
AC P50426;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE N-acetylglucosamine-6-sulfatase precursor (EC 3.1.6.14) (G6S)
DE (Glucosamine-6-sulfatase).
GN GNS.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95329494; PubMed=7605804;
RA Frederici K., Cavanagh K.T., Leipprandt J.R., Traviss C.E.,
RA Anson D.S., Hopwood J.J., Jones M.Z.;
RT "Cloning and sequence analysis of caprine N-acetylglucosamine 6-
RT sulfatase cDNA.";
RL Biochim. Biophys. Acta 1271:369-373 (1995).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate group of the N-
CC acetyl-D-glucosamine 6-sulfate units of heparan sulfate and
CC keratan sulfate.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- DISEASE: DEFECTS IN GNS RESULT IN THE ACCUMULATION OF PARTIALLY
CC DEGRADED HEPARAN SULFATE IN LYOSOMES CAUSING ORGANELLE, CELL AND
CC TISSUE DISTORTION, ULTIMATELY LEADING TO THE LYOSOMAL STORAGE
CC DISORDER, MUCOPOLYSACCHARIDOSIS TYPE IIID (MPS-IIID) (ALSO KNOWN
CC AS SANFILIPPO D SYNDROME).
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@sib-sib.ch).
EMBL; U17694; AB01663.1; --
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE 1; 1.
DR PROSITE; PS00149; SULFATASE 2; 1.
KW Hydrolase; Glycoprotein; Lysosome; Signal; Mucopolysaccharidosis.
FT SIGNAL 1 50
FT CHAIN 51 559
FT MOD_RES 98 98
FT N-ACETYLGLUCOSAMINE-6-SULFATASE.
FT 2-AMINO-3-OXOPROPIONIC ACID (BY
FT SIMILARITY).
FT CARBOHYD 118 118
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 124 124
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 190 190
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 205 205
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 324 324
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 369 369
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 412 412
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429
FT N-LINKED (GLCNAC. .) (POTENTIAL).

T CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
Q SEQUENCE 559 AA; 62711 MW; 786CCDC48334A458 CRC64;
Query Match 18.0%; Score 856; DB 1; Length 559;
Best Local Similarity 27.5%; Pred. No. 2.18-54;
Matches 231; Conservative 77; Mismatches 177; Indels 354; Gaps 20;
Y 3 PPSVLCLLSATVFLSGSSAFSLSHRLKGRFORDRNRIRPNILVLTDQDQVELGSMQ 62
b 29 PPLPLLL-LLGCLG-----VSGAAGSR---RPNVVLVLAADQDEVLGMT 72
Y 63 VNKTTRIMEQGGAHFNATVTPMCCPSRSILTGKIVHNHNTYTN--NENCSSPSHQA 120
b 73 PLKTKALIGEMGMTFSSAYVPSALCCPSRASILTGKYPHNHVNNTLEGNCSKSWQK 132
Y 121 QHESRTEAVLYNS--TGRTAFPGKLYNNG-----SYVPPGKWEVGLLKNRFRNYT 173
b 133 IQEPNTFFAILRSMCGYQTFAGKLYNBYGAPDAGGLGHVPLGHSYVALEKSKYNYT 192
Y 174 LCRNGVKEKHGSDYSKDYLTDLITNDSVSP--FRTSKMYPHRPVLMVISHAAPHGSDS 231
b 193 LSLNGKARKHGENYSDVLTDLVLANVSLDFLYKNSN-----PFFMMISTPAPHSPTA 247
Y 232 APOYSELFPNASQHTPSVYVAPNDKHMWRYT--GPMKPIHMBFTNMQEKLQTLMSV 290
b 248 APOYQAFNCFVAPRNKFN--IHGTNKHILIRQAFTMTNSIOFLNAPFRWQTLTSLV 306
Y 291 DDMETIYNMLVETGELONTIVYVYADHYHIGFGLVKGSMFYEDIRVFFVVRGNV 350
b 307 DDLVEKLVKLEFNGELNNTVYFTSDNGVHTGQPSLPIDKRQLYEDIDKVLPLVRGFI 366
Y 351 EAGLNPVHVLNIDAPLIDTILDIAGLDI--PADMDGSKILKILDTFRPNRFLKXKRVWR 409
b 367 KPNQTSKMLVANIDLPGLTIDTAGYLNKTDGMSFLPIL---RGASNL-----TWR 416
Y 410 DFLYVERGKLLHKRNDKVDAQEENFLPKYQKVDLCQAEYQYQACEQLGQKQCVEDAT 469
b 417 SDVLVE----- 422
Y 470 GKLLHKCKGPMRLGGSRALSNLVPKYVGQSEACTDSGDYKLSLAGRRKKLPKKYKA 529
b 423 -----YQEG----- 427
Y 530 SYVRSRSIRVALEVDGRVYHVLGDAQAQPNLTKRHWPGAPEDQDDKDGDFSGTGLP 589
b 428 -----RNV----- 431
Y 590 DYSAANPIKVRHCYLENTVQCDLILYKSLQAWKHLHIDHIEYTLQNKIKNLREVR 649
b 432 ----- 431
Y 650 GHLKKRPECDCXKISYTHQKGLKXGSSLHPRKGLQEKQKVLRLRQKXKXKLK 709
b 432 -----DPTCPSL----- 439
Y 710 LLKRLQNDTCSMPGLT--CFTHDNQHWOTAPFWTLGPFCACTSANNTYWCWRTINETHN 768
b 440 -----PGVSCF-----PDCVCEDAYNTYACVRTMSELWN 470
Y 769 FLFCPEF--ATGFLYEDLNTDYPQLMAVNTLDRDVLNQLHVLMLRSCGKYKOCNPR 825
b 471 LOYCEFDQDEVEFVYVNTADPHQLNNAKSIDPELLGKNYRLMWLQSCSGFTCTRTPR 529
RESULT 3
ID ABS VOLCA STANDARD; PRT; 649 AA.
AC Q10723;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).

OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94222055; PubMed=8168504;
RA Hallmann A., Sumper M.;
RT "An inducible arylsulfatase of Volvox carteri with properties
RT suitable for a reporter-gene system. Purification, characterization
RT and molecular cloning";
RL Eur. J. Biochem. 221:143-150(1994).
RN [2]
RP SEQUENCE OF 64-76, AND POST-TRANSLATIONAL MODIFICATION OF CYG-72.
RX MEDLINE=96283826; PubMed=8681943;
RA Selmer T., Hallmann A., Schmidt B., Sumper M., von Figura K.;
RT "The evolutionary conservation of a novel protein modification, the
RT conversion of cysteine to serinesemialdehyde in arylsulfatase from
RT Volvox carteri";
RL Eur. J. Biochem. 238:341-345(1996).
CC -I- FUNCTION: Is commonly produced by soil microorganisms and plays an
CC important role in the mineralization of sulfates.
CC -I- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -I- ENZYME REGULATION: INHIBITED BY NA(3)BO(3) AND KCN. NO INHIBITION
CC BY SODIUM DODECYL SULFATE, EVEN AT HIGH CONCENTRATION.
CC -I- SUBCELLULAR LOCATION: Periplasmic.
CC -I- INDUCTION: By sulfur deprivation.
CC -I- MISCELLANEOUS: The enzyme is thermostable, exhibiting a
CC temperature optimum at 60 degrees Celsius. Its optimal pH is 8.0.
CC -I- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; X77214; CAA54426.1; -.
DR PIR; S43229; S43229.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Periplasmic.
FT SIGNAL 1 22
FT CHAIN 23 649 ARYLSULFATASE.
FT MOD RES 72 72 2-AMINO-3-OXOPROPIONIC ACID.
FT DOMAIN 571 582 POLY-ALA.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 649 AA; 72287 MW; 0C23EFD77C43F7B9 CRC64;
Query Match 9.5%; Score 452.5; DB 1; Length 649;
Best Local Similarity 27.4%; Pred. No. 4.6e-25;
Matches 149; Conservative 82; Mismatches 170; Indels 143; Gaps 24;
QY 6 LVLCLLSATVFLSGSSAFSLSHRLKGRFORDRNRIRPNILVLTDQDQVELGSM--QV 63
b 7 VALCLLG---FAALTAAA---HQ-----RNFVVFVITDDQDQGNSTHPRY 47
QY 64 MNKTRIMEQGGAHFNATVTPMCCPSRSILTGKIVHNHNTYTNENCSSP-----SW 118
b 48 QPKLHEHRIYFGLHKLKNYFVTTTPVCCPSRLNWRGQFSN-----TNFTDVLGPHGYSKW 103
QY 119 QAQHSRTF--AVYLNSTGYRTAFPGKLYNLY-----NGSYVPPGKWEVGLLKNRFR--YNYT 173


```
Db 104 KSLGIDKSVLPWQLNGYNTYVVGKFLVDYSYQNVPAQWTDIDALVTPYTDYNNP 163
2y 174 -LCRNGVKEKGGSDYKDYLTDLITNDSVSPRTSKMYPHRPVLMWISHAAPHGPD 232
Db 164 GFSRNGATP---NIYPGFYSTVDIAKVAQAUKTA--VAAGKFFYAQISPIAPH- 212
2y 233 PQYSRLFPNASOHITPSYNYAPNP-DKHWMRYTGP- 275
Db 213 -TSTQIVFDVANATKTFYPPAPAPRHVELFSDATLPEGTSHKNLYEADVSDKAWIRA 271
2y 276 TNNLQ-----RKRLQTLMSVDDSMETINMLVETGELDNITYVYVYADHGHHGQ 324
Db 272 LPLAQNNRTYLEEVLRLRSASVDELIDRVVATLQEGVLDNTYLYISADNGVHVGT 331
2y 325 FGLVKGSMSEYEDIRVPFVVRGPNVEAG---CLNPHVILNIDLAFTILDLAG- 374
Db 332 HRFCAKGVATYDEDLRVFPFLIRGPGIRASHSDKPNASKVLHVDPAFTILTLAGAGDQVG 391
2y 375 -----LDIPADMKGSKILKLDTERPVPNRPHLKKRWVRDPLVE-----RKLHLH 421
Db 392 DKALDGTPLGLYANDGN---LLADYPRPANH-----RNQPGFVGWGSDEVILH 438
2y 422 KRNDKVDAAEENFLPKYO-----RVKDLQCORAEVQACEOLGQKQWCVEDATGKML 474
Db 439 -----HIPRYTNSWKAVRYD-----EDNQAWKLIYVCTNEREL 474
2y 475 HKCK 478
Db 475 YDLK 478

RESULT 4
ARS-CHLRE STANDARD; PRT; 647 AA.
AC P.4217;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
GN AS.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-55.
RC STRAIN=cw15;
RX MEDLINE=9384447; PubMed=2476654;
RA de Hostos E.L., Schilling J., Grossman A.R.;
RT "Structure and expression of the gene encoding the periplasmic
EL arylsulfatase of Chlamydomonas reinhardtii.";
RN Mol. Gen. Genet. 218:229-239(1989).
RP [2]
RP IDENTIFICATION OF PROBABLE FRAMESHIFTS.
RA Bauroch A.;
RL Unpublished observations (AUG-1996).
CC -1- CATALYTIC ACTIVITY: A phenol sulfite + H(2)O = a phenol + sulfate.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION. By sulfur deprivation.
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION. FRAMESHIFTS HAD TO BE
CC INTRODUCED TO MAXIMIZE THE SIMILARITY WITH OTHER SULFATASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16180; CAA34302.1; ALT_FRAME.
```

```
DR EMBL; X16179; CAA34301.1; ALT_FRAME.
DR EMBL; X52304; CAA36545.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Periplasmic.
FT SIGNAL 1 21
FT CHAIN 22 647
FT MOD_RES 73 73
FT 73
FT CARBOHYD 42 42
FT CARBOHYD 90 90
FT CARBOHYD 279 279
FT CARBOHYD 444 444
FT CARBOHYD 528 528
SQ SEQUENCE 647 AA; 72102 MW; 7404EAB1E233F0B1 CRC64;

Query Match 8.5%; Score 405; DB 1; Length 647;
Best Local Similarity 26.7%; Pred. No. 1.3e-21;
Matches 132; Conservative 76; Mismatches 179; Indels 108; Gaps 20;

QY 43 RPNILVLTDQDVELGSM--QVMNKTTRIMEQGAHFINAFVTTMCCPSRSILTKY 100
Db 26 KPNFVWIFTDDQDAIQNSTHPIYMPSLHKYIRYPGVVELSQYFVTTVPCCPSRTNLARGQF 85
QY 101 VHNHTVYNNENSSP-----SWQAQHESETP-AVILNSTGYETAFFGKYLNEY--NGS 151
Db 86 AHN-----TNFTSVLPFGYGWAKWGLGIDQSYLPDLKXQGYNTYVVGKFLVDYSVNYQ 141
QY 152 YVP-PGWKEWVGLLKNSRFYNTLCRNGVKRSGSDYSKDYLTDLITNDSVSPFRSKKM 210
Db 142 QVPRAGTISMEXVTPYTFDYNTRLQRNGATP---NIVPGEYSTDVIRDKGVAQIKSA--V 196
QY 211 YPHRPVLWVSHAHPGPDSPAQYSLFPNASHITPSSYNYAPNP-DKHWMRYTGP- 268
Db 197 AAGKFFYAQISPIAPH-----TSTQISTNPATGVTSTRSYFFPIAPPWHQVLFSDANLP 249
QY 269 -----KEI-----HMEFTNMQKRLQTLMSVDDSMETIYNMLV 302
Db 250 GGSXKNLYEVDVSDKPAWIRALPLAQNNRTQBEIYRLRLSL-GPDELIQVVKVTL 308
QY 303 ETGELDNTYVYADHGHHGQGLVKGSMPEFDIRVPVYVGRPNVEAG-----CLNPH 358
Db 309 EAGVLDNTYIISADNGYHVGARFGAGKTGYEEDLRVFLIRGPGIKASKSDXPQNSK 368
QY 359 IVLNIDLAFTILDLAG-----LDIPADMKGSKILKLDTERPVPNRPHLKKX 404
Db 369 VGLHVDPAFTILSLAGASHILGDKLDTGLYANDDG-----TLPSDYPRPEHQHQF 423
QY 405 MRV-----WRDSFLVERGKLLHKRDNKVDQAENFLPKYQVVKDLQORABYQTACEOLGQ 460
Db 424 QGEFVGWGSDELL-----QNLRSQPNNTW-KVVRTYD-----ESSKQ 459
QY 461 KWQCVEDATGKLLH 475
Db 460 GWKLIQAQCTNERELY 474

RESULT 5
VIDJ ECOLI
ID VIDJ ECOLI STANDARD; PRT; 497 AA.
AC P31447;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative sulfatase yidJ (EC 3.1.6.-).
GN YIDJ OR B3678.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
```

```

P SEQUENCE FROM N.A.
C STRAIN=K12 / MG1655;
X MEDLINE=93315143; PubMed=7686882;
C Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
T "DNA sequence and analysis of 136 kilobases of the Escherichia coli
T genome: organizational symmetry around the origin of replication.";
L Genomics 16:551-561(1993).
C -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
C -----
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C -----
R EMBL; L10328; AA62030.1; -.
R EMBL; AE000445; AAC76701.1; -.
R PIR; G65169; G65169.
R EcoGene; EG11705; YidJ.
R InterPro; IPR000917; Sulfatase.
R Pfam; PF00884; Sulfatase; 1.
R PROSITE; PS00523; SULFATASE_1; 1.
R PROSITE; PS00149; SULFATASE_2; 1.
W Hypothetical protein; Hydrolase; Complete proteome.
T ACT SITE 102 102 POTENTIAL.
I SEQUENCE 497 AA; 57295 MW; 2051C34EACEBA412E CRC64;
Query Match 6.4%; Score 302.5; DB 1; Length 497;
Best Local Similarity 25.5%; Pred. No. 2.4e-14;
Matches 105; Conservative 72; Mismatches 154; Indels 81; Gaps 21;
>Y 43 RPNILVLDDQDEL---GSMQVMKTR1-MEQGGAHFNAFVTPMCCPSRSILT 97
>b 3 RNFPLVMTDTQATNMVGCYSGKPLN-1QNIDSLAAEGIRFNSAVTCSFVCTPARAGLFT 61
>Y 98 GKXVHNHNTYNN---ENCSPSPQAQHESTFAVLNLTGYRTAFPGKYL---NEYNG 150
>b 62 GYANOSGPTNNAVAPGNIS---TMGRYFKDAGHYTCYIGKWHLDGHDIFG 110
>Y 151 S-YVPPGK---EWVGLLKNSRFNYTLCLNGVKRKGSDYSXDYLTDL-----ITN 198
>b 111 TGCPCPEWDADYFDCGANYLSLTETKEISLWNLGNSV--EDLOANHIDETFWAHRISN 168
>Y 199 DVSVPFRSKMYPHPVLMVLSHAAPCP-----EDSAPQYSRFLPNASCHITPSY 250
>b 169 RAVDFLQOPAR--ADEFFLMVSVYDEPHPTCPVEYLEKYADFYYELGKQAQDDL--- 222
>Y 251 NYAPNPKD--WIMRYTGM-----KEIHMEFTNMLQKQLQTLMSYDSDMETIYNML 301
>b 223 --ANKPEHRLNAQAMPSPVGGDGLYHPLYPACNDF-----VDDQIGRINAL 269
>Y 302 VERGELDNTYVYVADGHXIHGGFGLV-KGKSNPYEFDIRVPYVGRGNVEAGCLNPHIV 360
>b 270 TPE-QRENTWVIYTSDHGEMGAHLKISGAAM-YDDITRPLIIRSPQGERQVDTP-V 326
>Y 361 LNTDLAPTILDLADLDI PADMDGKSLTKLLDTERPVNRPHLKKKQKRVNEDSF 412
>b 327 SHIDLLPTMALADIEKPELLPGENILAKVEPRGVNVEFN---RYEIHDSF 375

```

RESULT 6

```

ID BETC_RHIME STANDARD; PRT; 512 AA.
AC O69787;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Choline-sulfatase (EC 3.1.6.6).
GN BETC OR R00949 OR SMC00127.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

```

```

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=102F34;
RX MEDLINE=97286552; PubMed=9141699;
RA Pocard J.A., Vincent N., Boncompagni E., Tombras Smith L.,
RA Poggi M.-C., Le Rudulier D.;
RT "Molecular characterization of the bet genes encoding glycine betaine
RT synthesis in Sinorhizobium meliloti 102F34.";
RL Microbiology 143:1369-1379(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaune V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=102F34;
RX MEDLINE=98409688; PubMed=9736747;
RA Oesteras M., Boncompagni E., Vincent N., Poggi M.-C., Le Rudulier D.;
RT "Presence of a gene encoding choline sulfatase in Sinorhizobium
RT meliloti bet operon: choline-O-sulfate is metabolized into glycine
RT betaine.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:11394-11399(1998).
CC -!- FUNCTION; CONVERTS CHOLINE-O-SULFATE TO CHOLINE.
CC -!- CATALYTIC ACTIVITY: Choline sulfate + H(2)O = choline + sulfate.
CC -!- PATHWAY: Osmoregulatory choline-glycine betaine pathway.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39940; AAC13371.1; -.
DR EMBL; AL591785; CAC45521.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Complete proteome.
FT CONFLICT 430 430
SQ SEQUENCE 512 AA; 58504 MW; ED7A4340258A6291 CRC64;
Query Match 5.5%; Score 260; DB 1; Length 512;
Best Local Similarity 22.5%; Pred. No. 3e-11;
Matches 92; Conservative 79; Mismatches 137; Indels 100; Gaps 19;
>Y 43 RPNILVLDDQDEL---GSMQVMKTR-RIMEQGAHFNAFVTPMCCPSRSILT 98
>b 5 RPNILITMVDQLNGKLPDGPADFLHAPNLKALAKRSARFHNNTSSPLCAPARAFMAG 64
>Y 99 KYVHNHNTYNNENCSPPSQQAQHESE--TFAYVLNSTGYRTAFPGK----- 143
>b 65 QLPSTRVYDN-----AAEVQSQSIPTYAHLRAGYTTALSGKHGFVGPDLHGFE 115
>Y 144 -----YINEYNGS--YVPPG-----KKEVGLLKNSRFNYTLCLNGVKRKGSDYSK 169
>b 116 ERLTDTYPADFGWTDPYRKPGRIDWYHNLGVSVGA-----GVAE----- 157
>Y 190 DYLTDLTINDSVSFRTSKQY-----PHRPVLMVLSHAAPHPEDSAPQYSRLFP 240
>b 190 DYLTDLTINDSVSFRTSKQY-----PHRPVLMVLSHAAPHPEDSAPQYSRLFP 240

```

158 --ITNQMEYDDEVAFLANQKLYQLSRENDESSRPWCLTVSFTPHDPYARRKFWDLIYE 215
 241 NASQHTPSYNYAP-----NPDKHWIMRYTGPMKPIHMFNTM-----LQKRB---LQTL 287
 216 DC-EHUTPEVGAIFDEQPHSQRM-----USCQYQNFVDTEENVRRSRAYFANI 266
 288 MSVDDSMETIYNMLVETGELDNITYVYTTADHGHIHQFGLVKGKSMPEYFDIRVFPYVRG 347
 267 SYLDEKVGELIDTLTRMLDDTLILFCSHDGMDLGERGL-WFKNMFEGSGARVPLMIAG 325
 348 PNVEACC-INPHVILNIDAPILDIAGL---DIPADMKGKILKILD 391
 326 PGIAPGLHLTP--TSNLDVTPTLADLAGISLEBVRPWTGDVSLVPMVN 371

RESULT 7

1D IDEN_ECOLI STANDARD; PRT; 560 AA.
 AC P77318; P78159;
 JT 16-OCT-2001 (Rel. 40, Created)
 JT 16-OCT-2001 (Rel. 40, Last sequence update)
 JT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative sulfatase yden precursor (EC 3.1.6.-).
 YDEN OR B1498.
 3S Escherichia coli.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12; MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.
 RA "The complete genome sequence of Escherichia coli K-12.";
 RT Science 277:1459-1474 (1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12;
 RC MEDLINE=97251357; PubMed=9097039;
 RX Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Moromura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.,
 RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377 (1996).
 CC 1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE000247; AAC74571.1; ALT INIT.
 DR EMBL; D90791; BAA15169.1; ALT INIT.
 DR EMBL; D90792; BAA15172.1; ALT_INIT.
 DR EcoGene; EG13796; yden
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase; 1.
 DR PROSITE; PS00523; SULFATASE_1; FALSE NEG.
 DR PROSITE; PS00149; SULFATASE_2; 1.
 KW Hypothetical protein; Hydrolase; Signal; Complete proteome.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 560 PUTATIVE SULFATASE YDEN.
 FT ACT_SITE 185 185 POTENTIAL.

SQ SEQUENCE 560 AA; 62802 MW; 67ED7FF7696C7A9F CRC64;
 Query Match 5.3%; Score 251; DB 1; Length 560;
 Best Local Similarity 22.7%; Pred. No. 1.5e-10;
 Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;
 QY 18 LLGGSFAFUSHHRLKGRFORDRNI-----RPMILVLTDDQ----- 54
 DB 18 LASGMAFAAHAAADDVKLKATKTNVAFSDFTTETXSTKGNPIIIVLTMDDLGYGQLPFDK 77
 QY 55 -----DVELGSMQVMKTRIMEQGAHFNAFVTTPMCPCPSR 93
 DB 78 GSFDPTKMNREVDYIKIDKAIKAAQKSTPILLSINDEGVRFNGVYAHGVSGPSRA 137
 QY 94 SILTKGYVHNHNTYNNENCSFSPWOAHE---SRTFAVYL-NSTGYRTAFPGK-VILNEY 148
 DB 138 AINTGRAPARFGVYSNTD-----AQDGIPLTETFLPELFQNHGYTAAVGNWLSKI 189
 QY 149 NGSYVPPGWKEWGLLKNSRFY--NYTL--CRNGVKEHGS DY----- 187
 DB 190 SNVPVPED-----KQTRDYHNDFTTSAEWEQPNRGFDYFMGFHAGTAYNPSL 241
 QY 188 -----SKDYLTDLITNDSVSFFRTSKMYPHREVLVMSHAAPHGPEDS-AP-QYSR 237
 DB 242 FKNRERVPAGYISDQLTDEAIGVDRAXTL--DQFFMLYLAYNAPHLPNENPAPDOYOK 299
 QY 238 LFNASOHITPSYNYAPNPKHWMRYTGMKPIEHMEFTNMLQRLQTLMSVDDSMETI 297
 DB 300 QFNTGSQ--TADNYA-----SVYSVDQGVKRI 325
 QY 298 YNMLVETGELDNITYVVTADHGVIH-GQF---GLVKG-KSMPEYFDIRVFPYVRGPNVEA 352
 DB 326 LEQLKNGQYDNTIILFTSDNGAVIDGLPLNGAQKGYKSQTYPGGTHTFMFMMW---K 381
 QY 353 GCLNP-----HIVNIDILAPILDIAGLIDPAD--MDGKSILKLLDTER---PVNRPHLKK 403
 DB 382 GKLPQGNKYDLISAMDYPTALDAADISIPKDLKLDGVSLLPWLDKQKQGEPHKNTWIT 441
 QY 404 KMRVWRDSFLVERGKLHK-----RNDN--KYDAQENFL 436
 DB 442 SYSHWFDEENIPFDWNVHKFRHQSDDYPHNPNTEDLSQFSYTVRRNDYSLVTVNNQL 501
 QY 437 PKYQVRKDLQRAEYQFACELQKQKQCV 465
 DB 502 GLY-KLTDLQCKNLAANPQVVKEMQGV 529
 ID STS HUMAN STANDARD; PRT; 583 AA.
 AC P08842;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Steryl-sulfatase precursor (EC 3.1.6.2) (Steroid sulfatase) (Steryl-
 DE sulfate sulfohydrolase) (Arylsulfatase C) (ASC).
 GN STS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RX [1]
 RX SEQUENCE FROM N.A.
 RX MEDLINE=89340479; PubMed=2668275;
 RA Stein C., Hille A., Seidel J., Rijnbout S., Waheed A., Schmidt B.,
 RA Geuze H., von Figura K.;
 RT "Cloning and expression of human steroid-sulfatase. Membrane
 RT topology, glycosylation, and subcellular distribution in BHK-21
 RT cells.";
 RL J. Biol. Chem. 264:13865-13872 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=87187642; PubMed=3032454;

Yen P.H., Allen E., Marsh B., Mohandas T., Wang N., Taggart R.T., Shapiro L.J.;
"Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X-Y interchange."; Cell 49:443-454(1987).
[3]
SEQUENCE OF 134-174 AND 461-583 FROM N.A.
MEDLINE=89077541; PubMed=3203382;
Yen P.H., Marsh B., Allen E., Tsai S.P., Ellison J., Connolly L., Neiswanger K., Shapiro L.J.;
"The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution."; Cell 55:1123-1135(1988).
[4]
SEQUENCE OF 22-45.
TISSUE=Liver;
MEDLINE=89352671; PubMed=2765556;
Kawano J.-I., Kotani T., Ohtaki S., Minamino N., Matsuo H., Oinuma T., Aikawa E.;
"Characterization of rat and human steroid sulfatases."; Biochim. Biophys. Acta 997:199-205(1989).
[5]
VARIANTS XLI LEU-341; ARG-372 AND TYR-446.
MEDLINE=92170784; PubMed=1539590;
Basler E., Grompe M., Parenti G., Yates J., Ballabio A.;
"Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis."; Am. J. Hum. Genet. 50:483-491(1992).
[6]
VARIANTS XLI LEU-341; ARG-372; SER-372; ARG-444 AND TYR-446.
MEDLINE=97400563; PubMed=9252398;
Alperin E.S., Shapiro L.J.;
"Characterization of point mutations in patients with X-linked ichthyosis. Effects on the structure and function of the steroid sulfatase protein."; J. Biol. Chem. 272:20756-20763(1997).
[7]
VARIANT XLI PRO-560.
MEDLINE=20146224; PubMed=10679952;
Sugawara T., Shimizu H., Hoshi N., Fujimoto Y., Nakajima A., Fujimoto S.;
"PCR diagnosis of X-linked ichthyosis: identification of a novel mutation (B560P) of the steroid sulfatase gene."; Hum. Mutat. 15:296-296(2000).
[8]
VARIANT XLI ARG-380.
MEDLINE=20304877; PubMed=10844566;
Oyama N., Satoh M., Iwatsuki K., Kaneko F.;
"Novel point mutations in the steroid sulfatase gene in patients with X-linked ichthyosis: transfection analysis using the mutated genes."; J. Invest. Dermatol. 114:1195-1199(2000).
-!- FUNCTION: CONVERSION OF SULFATED STEROID PRECURSORS TO ESTROGENS DURING PREGNANCY.
-!- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate + H₂O = 3-beta-hydroxyandrost-5-en-17-one + sulfate.
-!- SUBUNIT: Homodimer.
-!- SUBCELLULAR LOCATION: MICROSOMAL MEMBRANE. THE SEQUENCE SHOWS SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE PROTEIN IN THE MICROSOMAL MEMBRANE.
-!- DISEASE: Defects in STS are the cause of x-linked ichthyosis (XLI), a diskarization disorder characterized by the presence of prominent scales, especially on the neck, extremities, trunk, and buttocks. It affects approximately 1 of 2000-6000 males.
-!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

or send an email to license@isb-sib.ch.

CC EMBL; J04964; AAA60597.1; -
CC EMBL; M16505; AAA60596.1; -
DR EMBL; M23945; AAA60598.1; -
DR EMBL; M23567; AAA60599.1; -
DR PIR; A32641; KJHUAC.
DR HSP; P15848; IFSU.
DR Genew; HGNC:11425; STS.
DR MIM; 308100; -
DR GO; GO:0005783; C:Endoplasmic reticulum; TAS.
DR GO; GO:0005768; C:Endosome; TAS.
DR GO; GO:0005794; C:Golgi apparatus; TAS.
DR GO; GO:0005764; C:Lysosome; TAS.
DR GO; GO:0005624; C:Membrane fraction; TAS.
DR GO; GO:0005792; C:Microsome; TAS.
DR GO; GO:0005886; C:Plasma membrane; TAS.
DR GO; GO:0004773; P:steryl-sulfatase activity; TAS.
DR GO; GO:0008544; P:epidermal differentiation; TAS.
DR GO; GO:0006706; P:steroid catabolism; TAS.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;
KW Pregnancy; Signal; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 583
FT MOD_RES 75 75
FT STERYL-SULFATASE.
FT 2-AMINO-3-OXOPROPIONIC ACID (BY
FT SIMILARITY).
FT POTENTIAL.
FT POTENTIAL.
FT POTENTIAL.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT S -> L (IN XLI; LOSS OF ACTIVITY).
FT /FTID=VAR 007240.
FT W -> R (IN XLI; LOSS OF ACTIVITY).
FT /FTID=VAR 007241.
FT W -> S (IN XLI; LOSS OF ACTIVITY).
FT /FTID=VAR 014020.
FT G -> R (IN XLI).
FT /FTID=VAR 014021.
FT H -> R (IN XLI; LOSS OF ACTIVITY).
FT /FTID=VAR 014022.
FT C -> Y (IN XLI; LOSS OF ACTIVITY).
FT /FTID=VAR 007242.
FT Q -> P (IN XLI).
FT /FTID=VAR 014023.
FT A -> E (IN REF. 2).
FT CONFLICT 23 23
SQ SEQUENCE 583 AA; 65492 MW; 74746AFA9D21A0A6 CRC64;
Query Match 5.0%; Score 238.5; DB 1; Length 583;
Best Local Similarity 20.7%; Pred. No. 1.3e-09;
Matches 126; Conservative 67; Mismatches 180; Indels 237; Gaps 23;
QY 43 RPNILAVLTDODVELGSMQVM-NKTRRI-----MEQGAHAFINAFVTTPMCCPSRSSIL 96
Db RPNILVLMAD--DLGIGDPCYGNKTIPTPNDRLASGVGVKUTQHLAASPLCTPSRAAFM 83
QY 97 TGYVHNHTYNNENCSFSPWQ-----ACHESRTFAYLNSTGYRTAFFOK 143
Db TGRY-----PVRSGMASWSRTGVFLFTASSGGLPTDEITPAKLLKDGQSTALIGK 134
QY 144 Y-----LNEY-----NGSYVPPGNKE----- 159
Db WHLGMSCHSKTDFCHRHPLHGFNYFGISLTNLRCKPGEGSVFTTGFRLVFLQIVG 194
QY 160 -----WVGLLKNSRFNYVTLCRNGVKRKHG 184
Db VTLLTLAALNCILGLLHVPLGVFSLFLAALLTLTLFLGLFHYFRPLNCFMKN--YEIIQ 252

- SEQUENCE FROM N.A.
MEDLINE=94063929; PubMed=8244397;
Wilson P.J., Meaney C.A., Hopwood J.J., Morris C.P.;
"Sequence of the human iduronate 2-sulfatase (IDS) gene.";
Genomics 17:773-775(1993).
[3]
SEQUENCE FROM N.A.
MEDLINE=96352905; PubMed=8717057;
Timms K.M., Lu F., Shen Y., Pierson C.A., Muzny D.M., Gu Y.,
A Nelson D.L., Gibbs R.A.;
"130 kb of DNA sequence reveals two new genes and a regional
duplication distal to the human iduronate-2-sulfate sulfatase
locus.";
Genome Res. 5:71-78(1995).
[4]
SEQUENCE FROM N.A.
Muzny D., Ansari-Lari M.A., Timms K.M., Yu W., Dugan S., Lu J.,
Shen Y., Rowland K., Liu W., Perez L., Ding Y., Gonzalez O.,
Haywood M., Jain A., Leal B., Logan O., Nguyen V., Savage L., Shen H.,
Worley K., Chen E., Forcum J., Aronson A.D., Chiu M.W., Gorrell J.H.,
Brundage E., Di W., Chinault C., Nelson D., Gibbs R.A.;
Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A. (ISOFORM SHORT).
TISSUE=Lymphocytes;
MEDLINE=96079126; PubMed=8530090;
Malgren H., Carlberg B.M., Pettersson U., Bondeson M.L.;
"Identification of an alternative transcript from the human iduronate-
2-sulfatase (IDS) gene.";
Genomics 29:291-293(1995).
[6]
SEQUENCE OF 1-398 FROM N.A.
MEDLINE=93258349; PubMed=8490623;
Flomen R.H., Green E.P., Green P.M., Bentley D.R., Giannelli F.;
"Determination of the organisation of coding sequences within the
iduronate sulphate sulphatase (IDS) gene.";
Hum. Mol. Genet. 2:5-10(1993).
[7]
REVIEW ON MPS-II VARIANTS.
MEDLINE=94154729; PubMed=8111411;
Hopwood J.J., Bunge S., Morris C.P., Wilson P.J., Steglich C.,
Beck M., Schwinger E., Gal A.;
"Molecular basis of mucopolysaccharidosis type II: mutations in the
iduronate-2-sulphatase gene.";
Hum. Mutat. 2:435-442(1993).
[8]
VARIANTS MPS-II ARG-135 AND GLY-422.
MEDLINE=93265059; PubMed=1303211;
Hopwood J.J., Gal A.;
"Mutation analysis of the iduronate-2-sulfatase gene in patients with
mucopolysaccharidosis type II (Hunter syndrome).";
Hum. Mol. Genet. 1:335-339(1992).
[9]
VARIANT MPS-II TRP-468.
MEDLINE=93258320; PubMed=1284597;
Crotti P.L., Bunge S., Anderson R.A., Whitley C.B.;
"Mutation R468W of the iduronate-2-sulfatase gene in mild Hunter
syndrome (mucopolysaccharidosis type II) confirmed by in vitro
mutagenesis and expression.";
Hum. Mol. Genet. 1:755-757(1992).
[10]
VARIANTS MPS-II ARG-86; ASP-94; ARG-120; PRO-221 AND GLY-422.
MEDLINE=94108441; PubMed=8281149;
Bunge S., Steglich C., Zuther C., Beck M., Morris C.P., Schwinger E.,
Schinzel A., Hopwood J.J., Gal A.;
"Iduronate-2-sulfatase gene mutations in 16 patients with
mucopolysaccharidosis type II (Hunter syndrome).";
Hum. Mol. Genet. 2:1871-1875(1993).
[11]
VARIANTS MPS-II GLU-68; HIS-293; GLY-478 AND ARG-485.
MEDLINE=95072615; PubMed=7981716;
Schroeder W., Wulff K., Wehnert M., Seidlitz G., Herrmann F.H.;
"Mutations of the iduronate-2-sulfatase (IDS) gene in patients with
Hunter syndrome (mucopolysaccharidosis II).";
Hum. Mutat. 4:128-131(1994).
[12]
VARIANTS MPS-II TRP-132; TYR-239; ARG-358; HIS-469 AND CYS-523.
MEDLINE=95193786; PubMed=7887413;
Jonsson J.J., Aronovich E.L., Braun S.E., Whitley C.B.;
"Molecular diagnosis of mucopolysaccharidosis type II (Hunter
syndrome) by automated sequencing and computer-assisted
interpretation: toward mutation mapping of the iduronate-2-sulfatase
gene.";
Am. J. Hum. Genet. 56:597-607(1995).
[13]
VARIANTS MPS-II LEU-86; ASN-87; PRO-92; ASN-135; CYS-345 AND TRP-468.
MEDLINE=95245347; PubMed=7728156;
Popowska E., Rathmann M., Tytki-Szymanska A., Bunge S., Steglich C.,
Schwinger E., Gal A.;
"Mutations of the iduronate-2-sulfatase gene in 12 Polish patients
with mucopolysaccharidosis type II (Hunter syndrome).";
Hum. Mutat. 5:97-100(1995).
[14]
VARIANT MPS-II VAL-346.
MEDLINE=95322987; PubMed=7599640;
Li P., Huffman P., Thompson J.N.;
"Mutations of the iduronate-2-sulfatase gene on a T146T background in
three patients with Hunter syndrome.";
Hum. Mutat. 5:272-274(1995).
[15]
VARIANTS MPS-II.
MEDLINE=97094177; PubMed=8940265;
Rathmann M., Bunge S., Beck M., Kresse H., Tytki-Szymanska A., Gal A.;
"Mucopolysaccharidosis type II (Hunter syndrome): mutation 'hot spots'
in the iduronate-2-sulfatase gene.";
Am. J. Hum. Genet. 59:1202-1209(1996).
[16]
VARIANTS MPS-II LEU-333 AND ASP-346.
MEDLINE=96163494; PubMed=8566953;
Olsen T.C., Eiken H.G., Knappskog P.M., Kase B.F., Mansson J.-E.,
Bowman H., Apold J.;
"Mutations in the iduronate-2-sulfatase gene in five Norwegians with
Hunter syndrome.";
Hum. Genet. 97:198-203(1996).
[17]
VARIANTS MPS-II LEU-333 AND GLY-334.
MEDLINE=96255081; PubMed=8830188;
Li P., Thompson J.N.;
"Detection of four novel mutations in the iduronate-2-sulphatase gene
by single-strand conformation polymorphism analysis of genomic
amplicons.";
J. Inherit. Metab. Dis. 19:93-94(1996).
[18]
VARIANTS MPS-II ASP-63; THR-347; GLN-468 AND LEU-468.
MEDLINE=97365936; PubMed=9222763;
Vallani G.R.D., Balzano N., Grosso M., Salvatore F., Izzo P.,
di Natale P.;
"Mucopolysaccharidosis type II: identification of six novel mutations
in Italian patients.";
Hum. Mutat. 10:71-75(1997).
[19]
VARIANT MPS-II GLN-468.
MEDLINE=98041699; PubMed=9375851;
Sukegawa K., Song X.-Q., Masuno M., Fukao T., Shimozawa N., Fukuda S.,
Isogai K., Nishio H., Matsuo M., Tomatsu S., Kondo N., Orii T.;
"Hunter disease in a girl caused by R468Q mutation in the
iduronate-2-sulfatase gene and skewed inactivation of the X
chromosome carrying the normal allele.";
Hum. Mutat. 10:361-367(1997).
[20]
VARIANTS MPS-II ASN-45; TYR-115; LEU-228; ARG-266; LYS-434; LYS-485
AND CYS-502.
MEDLINE=99092178; PubMed=9875019;
Vafiadaki E., Cooper A., Heptinstall L.E., Hatton C.E., Thornley M.,
Wraith J.E.;

"Mutation analysis in 57 unrelated patients with MPS II.";
Arch. Dis. Child. 79:237-241(1998).

[21] VARIANTS MPS-II.
MEDLINE=9211886; PubMed=921913;
Karsten S., Voskoboeva E., Tishkagina S., Pettersson U.,
Krasnopol'skaja X., Bondeson M.-L.;
"Mutational spectrum of the iduronate-2-sulfatase (IDS) gene in 36
unrelated Russian MPS II patients.";
Hum. Genet. 103:732-735(1998).

[22] VARIANTS MPS-II LEU-86; HIS-88; PRO-88; ILE-118 AND HIS-266.
Balzano N., Villani G.R.D., Grosso M., Izzo P., di Natale P.;
"Detection of four novel mutations in the iduronate-2-sulfatase
gene.";
Hum. Mutat. 11:333-333(1998).

[23] VARIANTS MPS-II THR-85; HIS-88; ILE-349 AND VAL-521.
MEDLINE=98112423; PubMed=9452044;
Gott L., Coll M.O., Chabas A.;
"Mutations in the iduronate-2-sulfatase gene in 12 Spanish patients
with Hunter disease.";
Hum. Mutat. Suppl. 1:S66-S68(1998).

[24] VARIANTS MPS-II PHE-143; TRP-184; VAL-269 AND HIS-348.
Karsten S.L., Vokkova E., Carlberg B.-M., Kleijer W.J.,
Toennesen I., Petersson U., Bendesen M.-L.;
"Identification of 9 novel gene mutations in 19 unrelated Hunter
syndrome (Mucopolysaccharidosis type II) patients";
Hum. Mutat. 12:433-433(1998).

[25] VARIANTS MPS-II PHE-143; LYS-341; TYR-342 AND PHE-491. MEDLINE=99235558; PubMed=10220152;

every Match 4.9%; Score 234.5; DB 1; Length 550;
 t Local Similarity 21.2%; Pred. No. 2.3e-09;
 Conserves 114; Conservative 80; Mismatches 189;
 Indels 155; Gaps 23;

```

1  MGPPS-----LVLCILSATVPSLLGGSAFLSHRLKGRFQDRDRNIRNIIILVLTDDOD 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1  MPFPRGRLGLLWGLVLSVCVALGSET-----QANSTTDAUNVLIIIVDDLR 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

56  VELGSOVMNKTRRIWEGGAH-----FINAVTTPMCCPSRSSILTK-----YVHNHTY 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49  PSLGCGDKLVRSPINDOLASHSLLFQNAFAQACVAPSEVSFLTGRRTDTRLVDNFYS 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

108  TNNENCSPSWAOAHES-RTFAVILNSTGYRTAFPGKYL-----NEYNGSVYP- 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109  -----WRVHAGNSTTPOYPKENGIVYMSVGKVPFGPISSNHTDDSPYSNSPPY 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

155  -PGKWEVGLLNKNSRPNYNTLCRNGVKEKGS-----DYSKOYLT-D-LTNDOSVFF 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159  HP-----SSEYENTKTCRPGDGLHANLLCPVDVLDPVEGTLPPQKQTEQAIQLL 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

205  RTSKQVYHRPVLWVISHAAPHGSDSPAQYSRLFFNASOHITPSYVNAPNPD----- 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210  EKOKT--SASPFLAVGVHKPHIFRYPK3QKLYP--LENIT---LAPDEVDPGLPP 261
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

258  ---XZHI-----WRYTGMKPIHMETNMLQRKLOTLMSVDDSNETIYNKLVEY 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
262  VAYNPMWDIRQEDVOALNISVPYGFIPVDSPQRKIOSYPASVSLDTQVGRLLSALDDL 321
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

305  GELDNTIYVTADHGHHGQGLVKGKSMPEYFDIRP--FYVRGN----- 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
322  QLANSTIIAFTSDHGVALGEHG--EWAKYSNPDVATHVELIFYVPGRTASLPEAGEKLFPY 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

350  -----VZAGCLNPHIHLNIDLAPTILDLIAGLDI PADMDGKILKLLOTEPVNRPF 399
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
381  LDPFDSASQLMDFGQSDMLVELVSLFTLAGLAGLQVP-----PRCPVPSPF 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

400  HLKKOMRVWRD3FLYVERGKLLKLRDNKDVAQEEFNFLP-----KYORVKOLCQ 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428  HVE-----LCREGNLLKHFRPF-DUEEDPYLFGNPRELIAYSOVPPSPDIPO 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


QY 94 SILTGKVVHNTYNNENCSWSQ---AQHESRTFAVYLNSTGYRTAFKGY---LNE 147
 DB 106 AIMGRLPVRIIGF--GETRVFLEWTKTGLPKSELTIAEAMKEAGVATGVMGKWLHGINE 163
 QY 148 ---YNGSYVP-----PGWKEW-----VGLLK-----NSRFY---NYTLRNGVKE 181
 DB 164 NSSTDGAHFNHGFDFVGHNLFTFNSWSDDTGLHKKDFPDSQRCVLYNATLV----- 217
 QY 182 KHGSDYSKDYLTDLITNDSVSFFRTSKMYPHVPVLMVISHAAPHGPDSPAPQYSLPEN 241
 DB 218 --SQYQHKGLTQTFDDALGFI-----EDN----- 241
 QY 242 ASQHTPSYNYAPNPKHWMRYTGMKPHHMFETNMLQKRL--QTLMSVDSMETIYNM 300
 DB 242 ---HADPFYFAVAFHMTLSFSSD-----DFSCSTRGRYCDNLEHGDVAKIVDK 291
 QY 301 LVTEGELDNVYIVYADHGVHI-----GQFGLVK-GKSNPFEDIRVFPVYRGNVPEAG 353
 DB 292 LEENISENTIIIFISDHGPHRYCEGGDASIFRGKSHSWEGGHRIPYIVYWPQTISP 351
 QY 354 CLNPHVINLIDLAFTIIDLADIPAD--MDGKSILKLL--DTERPVNRFHLKMKRVR 409
 DB 352 GISNEIVTSMIITATRADLGTTLPDRIYDGKSIKDVLEGSASPHSFFYYCK----- 406
 QY 410 DSFLVER-GKLLHCRNDKVDAGEE 433
 DB 407 DNLMAVRVGKYKAHFTQVRVSODE 431

RESULT 12
 IDS_MOUSE
 ID IDS_MOUSE STANDARD; PRT; 563 AA.
 AC Q08890;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Iduronate 2-sulfatase precursor (EC 3.1.6.13).
 GN I2S.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RX MEDLINE=93315172; PubMed=9325651;
 RA Daniele A., Faust C.J., Herman G.E., di Natale P., Ballabio A.;
 RT "Cloning and characterization of the cDNA for the murine iduronate
 RL sulfatase gene";
 RL Genomics 16:755-757(1993).
 CC -1- FUNCTION: REQUIRED FOR THE LYSOSOMAL DEGRADATION OF HEPARAN
 CC -1- SULFATE AND DERMATAN SULFATE.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 2-sulfate groups of the L-
 CC iduronate 2-sulfate units of dermatan sulfate, heparan sulfate and
 CC heparin.
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L07921; AAA37880.1; -
 CC PIR; A47153; A47153.
 CC MGI; 96417; IGS.
 CC InterPro; IPR000917; Sulfatase.
 CC Pfam; PF00884; Sulfatase; 1.
 CC PROSITE; PS00523; SULFATASE_1; 1.

DR PROSITE; PS00149; SULFATASE 2; 1.
 KW Hydrolase; Glycoprotein; Lysozyme; Zymogen; Signal.
 FT SIGNAL ? POTENTIAL
 FT PROPEP ? 59 BY SIMILARITY
 FT CHAIN 60 563 IDURONATE 2-SULFATASE.
 FT MOD_RES 99 99 2-AMINO-3-OXOPROPIONIC ACID (BY
 FT SIMILARITY).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 563 AA; 63437 MW; 52D008F91BA88FAC CRC64;
 Query Match 4.8%; Score 226; DB 1; Length 563;
 Best Local Similarity 21.7%; Pred. No. 9.9e-09;
 Matches 101; Conservative 69; Mismatches 199; Indels 96; Gaps 17;
 QY 16 PSLLGSSAFUSHRLKGRFORRRNIRPMILVLTDQDVELGSMQVMNKRIMEQGG 75
 DB 28 PSLLGSCFIALESAAQNGSATDAL---NILLTIIVDDLRFSLGCGYDKLVRSNIDQLA 83
 QY 76 AF---FINAFVTTPMCCPSRSSIILTKG-----YVHNENTYNNENCSPPKQAQHE-RT 126
 DB 84 SHSVLFQFAQAQAVCAPSRVSLTGRRPDTRTLYDFNSY-----WRVHSGMFST 133
 QY 127 FAVYLNSTGYRTAFKGYL-----NEYNGSYVPVPGWKEVGLLKNRFFNYTLCRNGVKE 181
 DB 134 IPQYFKENGYYTMSVGKVFHFGISSNHSDDY-PYSWSFPPVHPSSEKYENTKTCKGQDGK 192
 QY 182 KHG-----SDYSKDYLTDLITNDSVSFFRTSKMYPH-RPVLVMIISHAAPHGPDSDA 232
 DB 193 LHANLCPVDVADYVPEGLTPDKQSTEEA--IRLEKMKTSGPSFFFLAVGVYKHPHIFRYP 250
 QY 233 POYSRLFPNASOHIPTPSYNYAPNPKH-----PDHPVDSLPPVAYNPWMDIREREDVQALNISVYGP 270
 DB 251 KEFQKLYP--LENITLA-----PDPHVDPSLPVAYNPWMDIREREDVQALNISVYGP 302
 QY 271 IHMEFTNMLQKRLQTLMSVDSMETIYNMLVETGELDNVYIVYADHGVHIGQFG----- 326
 DB 303 IPEDFORKIQSYFASVYLDYQVGHVLSALDDLRLAHNTIIAFTSDHGWALGEGHGEWAK 362
 QY 327 -----LVKGKSNPYE-----FDIRVFPVYRGNVPEAGCLNPHVINLIDLA 366
 DB 363 YSNFDVATRVPLMLYVFORATPLPAQCKLFPDRDPDPPASDMDWAGRTEDVELVSLF 422
 QY 367 PTILDIAGLDIPADMKGSIKLLDTERPVNRFHLKMKRVRWDS 411
 DB 423 PTLAGLAGL--PVLGAPSLFLMLSFPAEKARIFRASICSSMTWKS 465

RESULT 13
 IDS_MOUSE
 ID GAGS_HUMAN STANDARD; PRT; 522 AA.
 AC P34059;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE N-acetylgalactosamine-6-sulfatase precursor (EC 3.1.1.6.4) (N-
 DE acetylgalactosamine-6-sulfate sulfatase) (Galactose-6-sulfate
 DE sulfatase) (GalNAc6S sulfatase) (Chondroitinsulfatase)
 DE (Chondroitinase).
 GN GALNS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=92095973; PubMed=1755850;
 RA Tomatsu S., Fukuda S., Masue M., Sukegawa K., Fukao T., Yamagishi A.,

RA Hori T., Iwata H., Ogawa T., Nakashima Y., Hanyu Y., Hashimoto T.,
RA Titani K., Oyama R., Suzuki M., Yagi K., Hayashi Y., Orii T.,
RT "Morquio disease: isolation, characterization and expression of full-
RT length cDNA for human N-acetylgalactosamine-6-sulfate sulfatase.";
RL Biochem. Biophys. Res. Commun. 181:677-683(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95095267; PubMed=8001980;
RA Morris C.P., Guo X.H., Apostolou S., Hopwood J.J., Scott H.S.;
RT "Morquio A syndrome: cloning, sequence, and structure of the human N-
RT acetylgalactosamine-6-sulfatase (GALNS) gene.";
RL Genomics 22:652-654(1994).
RN [3]
RP VARIANT MPS-IVA LYS-204.
RX MEDLINE=92391522; PubMed=1522213;
RA Fukuda S., Tomatsu S., Masue M., Sukegawa K., Iwata H., Ogawa T.,
RA Nakashima Y., Hori T., Yamagishi A., Hanyu Y., Morooka K., Kinn T.,
RA Hashimoto T., Orii T.;
RT "Mucopolysaccharidosis type IVA. N-acetylgalactosamine-6-sulfate
RT sulfatase exonic point mutations in classical Morquio and mild
RT cases.";
RL J. Clin. Invest. 90:1049-1053(1992).
RN [4]
RP VARIANTS MPS-IVA.
RX MEDLINE=95397840; PubMed=7668283;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G.,
RA Yamagishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N.,
RA Suzuki Y., Shimozawa N., Orii T.;
RT "Mucopolysaccharidosis IVA: identification of a common missense
RT mutation 1113F in the N-Acetylgalactosamine-6-sulfate sulfatase
RT gene.";
RL Am. J. Hum. Genet. 57:556-563(1995).
RN [5]
RP VARIANTS MPS-IVA.
RX MEDLINE=95315929; PubMed=7795586;
RA Ogawa T., Tomatsu S., Fukuda S., Yamagishi A., Maruf Rezvi G.,
RA Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.;
RT "Mucopolysaccharidosis IVA: screening and identification of mutations
RT of the N-acetylgalactosamine-6-sulfate sulfatase gene.";
RL Hum. Mol. Genet. 4:341-349(1995).
RN [6]
RP VARIANTS MPS-IVA ARG-77; TRP-90; VAL-96; LEU-151; GLY-230 AND THR-291.
RX MEDLINE=95359983; PubMed=7633425;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G.,
RA Yamagishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N.,
RA Suzuki Y., Shimozawa N., Orii T.;
RT "Mucopolysaccharidosis type IVA: identification of six novel
RT mutations among non-Japanese patients.";
RL Hum. Mol. Genet. 4:741-743(1995).
RN [7]
RP VARIANT MPS-IVA SER-487.
RX MEDLINE=96047158; PubMed=7581409;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamada N., Isogai K.,
RA Kato Z., Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.;
RT "Two new mutations, Q473X and N487S, in a Caucasian patient with
RT mucopolysaccharidosis IVA (Morquio disease).";
RL Hum. Mutat. 6:195-196(1995).
RN [8]
RP VARIANTS MPS-IVA ALA-138; SER-151 AND LEU-151.
RX MEDLINE=96216146; PubMed=8651279;
RA Tomatsu S., Fukuda S., Yamagishi A., Cooper A., Wraith J.E., Hori T.,
RA Kato Z., Yamada N., Isogai K., Sukegawa K., Kondo N., Suzuki Y.,
RA Shimozawa N., Orii T.;
RT "Mucopolysaccharidosis IVA: four new exonic mutations in patients
RT with N-acetylgalactosamine-6-sulfate sulfatase deficiency.";
RL Am. J. Hum. Genet. 58:950-962(1996).
RN [9]
RP VARIANTS MPS-IVA CYS-94 AND VAL-97.
RX MEDLINE=96423834; PubMed=8826435;
RA Cole D.E.C., Fukuda S., Gordon B.A., Rip J.W., Lecouteur A.N.,
RA Rupar C.A., Tomatsu S., Ogawa T., Sukegawa K., Orii T.;
RT "Heteroallelic missense mutations of the galactosamine-6-sulfate
RT sulfatase (GALNS) gene in a mild form of Morquio disease (MPS IVA).";

RL Am. J. Med. Genet. 63:558-565(1996).
RN [10]
RP VARIANTS MPS-IVA.
RX MEDLINE=97442274; PubMed=9298823;
RA Bunge S., Kleijer W.J., Tytki-Szymanska A., Steglich C., Beck M.,
RA Tomatsu S., Fukuda S., Poorthuis B.J.H.M., Czartoryska B., Orii T.,
RA Gal A.;
RT "Identification of 31 novel mutations in the N-acetylgalactosamine-6-
RT sulfatase gene reveals excessive allelic heterogeneity among patients
RT with Morquio A syndrome.";
RL Hum. Mutat. 10:223-232(1997).
RN [11]
RP VARIANTS MPS-IVA.
RX MEDLINE=98041700; PubMed=9375852;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Ferreira P.,
RA di Natale P., Tortora P., Fujimoto A., Kato Z., Yamada N., Isogai K.,
RA Yamagishi A., Sukegawa K., Suzuki Y., Shimozawa N., Kondo N.,
RA Sly W.S., Orii T.;
RT "Fourteen novel mucopolysaccharidosis IVA producing mutations in
RT GALNS gene.";
RL Hum. Mutat. 10:368-375(1997).
RN [12]
RP VARIANTS MPS-IVA.
RX MEDLINE=98180718; PubMed=9521421;
RA Yamada N., Fukuda S., Tomatsu S., Muller V., Hopwood J.J., Nelson J.,
RA Kato Z., Yamagishi A., Sukegawa K., Kondo N., Orii T.;
RT "Molecular heterogeneity in mucopolysaccharidosis IVA in Australia
RT and Northern Ireland: nine novel mutations including T312S, a common
RT allele that confers a mild phenotype.";
RL Hum. Mutat. 11:202-208(1998).
RN [13]
RP VARIANTS SER-393 AND MET-488.
RX MEDLINE=98112415; PubMed=9452036;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamagishi A., Kato Z.,
RA Yamada N., Isogai K., Sukegawa K., Suzuki Y., Shimozawa N., Kondo N.,
RA Orii T.;
RT "Fifteen polymorphisms in the N-acetylgalactosamine-6-sulfate
RT sulfatase (GALNS) gene: diagnostic implications in Morquio disease.";
RL Hum. Mutat. Suppl. 1:S42-S46(1998).
RN [14]
RP CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate groups of the N-
RP acetyl-D-galactosamine 6-sulfate units of chondroitin sulfate and
RP of the D-galactose 6-sulfate units of keratan sulfate.
CC -!- SUBUNIT: OLIGOMER OF DISULFIDE LINKED 40- AND 15-KDA POLYPEPTIDES.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- DISEASE: Defects in GALNS are the cause of mucopolysaccharidosis
CC type IVA (MPS-IVA) [MIM:253000]; also known as Morquio A syndrome.
CC MPS-IVA is characterized by specific spondyloepiphyseal dysplasia,
CC short trunk dwarfism, coxa valga, odontoid hypoplasia, corneal
CC opacities, preservation of intelligence, and excessive urinary
CC excretion of keratan sulfate and chondroitin-6-sulfate. Severely
CC affected patients usually die of cardiopulmonary disturbance or
CC cervical cord compression in the second or third decade of life.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D17629; BAA04535.1; JOINED.
CC EMBL; D17616; BAA04535.1; JOINED.
CC EMBL; D17617; BAA04535.1; JOINED.
CC EMBL; D17618; BAA04535.1; JOINED.
CC EMBL; D17619; BAA04535.1; JOINED.
CC EMBL; D17620; BAA04535.1; JOINED.
CC EMBL; D17621; BAA04535.1; JOINED.
CC EMBL; D17622; BAA04535.1; JOINED.
CC EMBL; D17623; BAA04535.1; JOINED.
CC EMBL; D17624; BAA04535.1; JOINED.
CC EMBL; D17625; BAA04535.1; JOINED.

```

DR EMBL; D17626; BAA04535.1; JOINED.
DR EMBL; D17628; BAA04535.1; JOINED.
DR EMBL; U06088; AAC51350.1; JOINED.
DR EMBL; U06079; AAC51350.1; JOINED.
DR EMBL; U06078; AAC51350.1; JOINED.
DR EMBL; U06080; AAC51350.1; JOINED.
DR EMBL; U06081; AAC51350.1; JOINED.
DR EMBL; U06082; AAC51350.1; JOINED.
DR EMBL; U06083; AAC51350.1; JOINED.
DR EMBL; U06084; AAC51350.1; JOINED.
DR EMBL; U06085; AAC51350.1; JOINED.
DR EMBL; U06086; AAC51350.1; JOINED.
DR EMBL; U06087; AAC51350.1; JOINED.
DR PIR; JQ1299; KJHUG6.
DR HSP; P15848; IFSU.
DR GENE; HGNC:4122; GALNS.
DR MIM; 253000; -.
DR GO; GO:0003943; F:N-acetylgalactosamine-4-sulfatase activity; TAS.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Glycoprotein; Signal; Lysosome; Mucopolysaccharidosis;
KW Disease mutation; Polymorphism; Dwarfism.
FT SIGNAL 1 26
FT CHAIN 27 522 N-ACETYL GALACTOSAMINE-6-SULFATASE.
FT MOD_RES 79 79 2-AMINO-3-OXOPROPIONIC ACID (BY
FT SIMILARITY).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 47 47 G -> R (IN MPS-IVA; SEVERE FORM).
FT /FTID=VAR 007172.
FT VARIANT 60 60 D -> N (IN MPS-IVA; MILD FORM).
FT /FTID=VAR 007173.
FT VARIANT 67 67 L -> M (IN MPS-IVA; ASSOCIATED WITH S-409
Query Match 4.6%; Score 218.5; DB 1; Length 522;
Best Local Similarity 25.0%; Pred. No. 3.1e-08;
Matches 100; Conservative 54; Mismatches 151; Indels 95; Gaps 20;
2Y 44 PNILLVLTDDQ-VELG-----SMQVWNTKTRIMEQGAHFINAFVTTMCPSPSSILIT 97
31 PNILLVLTDDQ-VELG-----SMQVWNTKTRIMEQGAHFINAFVTTMCPSPSSILIT 97
98 GKTVHNTYTN--ENCSPPS---WQAHESRTFAVYLNSTGYRTAFPGKYLENYNGS 151
89 GLRPIRGFTTNAHARNAYTPQEIIVGIPDSQLPELLKAGYVSKIVGKWLGHHPQ 148
152 YVP--PWKEWUGLLKNSRPYNY-TLCRNGVK-----EKHGSDY-----SKDYLT 194
149 FHLKKGFDSEWG-SPNCHFGPYDNKARNPIPVYRDWEMVGRYEEFFINLKTGEANLTQ 207
195 LIITDSVSPRTSKMYPHRPVLMTVSHAAPHGPEDSAPQYSRLFPNASCHITPSVNYAP 254
208 IYLOEALDFI--KQARHHPFLYVAVDATHAPV-----YAS 242
255 NPKHWMIRVTGPMKPIHMEFTNMLQKEL-OFTMSVDSMETIYNNMLVETGEIDNTYIV 313
243 KP-----FLGTSORGRYGDVAVREIDDSIGKILELLQDLHVADNTFV 284
314 YTADHGWHI-----QGEF-LVKGKSMVEPDIRVPFYVGRP-NVEACLPNPHVIND 364
285 FTSNAGALISAPEQGSNGFPFCGQTTFEGWGRPALAMPGHVITAGQVSHQLGSIND 344
365 LAPTILDIAGLDIPAD--MDGKSL-----KLDDTERPV 396
345 LFTTSLALAGLTPPSDRAIDGLNLLPTLQGRLLMD--RRI 382
RESULT 14
ARSA MOUSE
ID ARSA_MOUSE STANDARD; PRT; 506 AA.
AC P50428;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Arylsulfatase A precursor (EC 3.1.6.8) (ASA) (Cerebroside-sulfatase).
GN ARSA OR AS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV, and C57BL/6J;
RX MEDLINE=94245194; PubMed=7910580;
RA Kreyling J., Polten A., Hess B., von Figura K., Menz K., Steiner F.,
RA Gieselmann V.;
RT "Structure of the mouse arylsulfatase A gene and cDNA.";
RL Genomics 19:249-256(1994).
RN [2]
RP SEQUENCE OF 32-66 FROM N.A.
RX MEDLINE=92241876; PubMed=1572648;
RA Grompe M., Pieretti M., Caskey C.T., Ballabio A.;
RT "The sulfatase gene family: cross-species PCR cloning using the MOPAC
RT technique.";
RL Genomics 12:755-760(1992).
CC -!- FUNCTION: Hydrolyzes cerebroside sulfate.
CC -!- CATALYTIC ACTIVITY: A cerebroside 3-sulfate + H(2)O = a
CC cerebroside + sulfate.
CC -!- SUBUNIT: EXISTS BOTH AS A SINGLE CHAIN OF 58 kDa (COMPONENT A)
CC OR AS A CHAIN OF 50 kDa (COMPONENT B) LINKED BY DISULFIDE BOND (S)
CC TO A 7 kDa CHAIN (COMPONENT C).
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X73231; CAA51703.1; -.
DR EMBL; X73230; CAA51702.1; -.
DR EMBL; M82876; AAA37260.1; -.
DR PIR; A54190; A54190.
DR HSP; P15289; LAUK.
DR MGP; MGI.88077; Arsa.
DR GO; GO:0000299; C: integral membrane protein of membrane fraction; IDA.
DR GO; GO:0005886; C: plasma membrane; IDA.
DR GO; GO:0007339; P: binding of sperm to zona pellucida; IMP.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Lysosome.
FT SIGNAL 1 17
FT CHAIN 18 506 2-AMINO-3-OXOPROPIONIC ACID (BY
FT MOD_RES 68 68 SIMILARITY).
FT DISULFID 155 171 BY SIMILARITY.
FT DISULFID 160 167 BY SIMILARITY.
FT DISULFID 299 413 BY SIMILARITY.
FT DISULFID 487 499 BY SIMILARITY.
FT DISULFID 488 501 BY SIMILARITY.
FT DISULFID 492 498 BY SIMILARITY.
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 506 AA; 53776 MW; 74BB16401CF18DF6 CRC64;
Query Match 4.4%; Score 210.5; DB 1; Length 506;
Best Local Similarity 25.2%; Pred. No. 1.1e-07;
Matches 94; Conservative 64; Mismatches 158; Indels 57; Gaps 18;

```

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J05225; AAAS1784.1; -
EMBL; M32373; AAAS1779.1; -
EMBL; X72735; CAAS1272.1; -
EMBL; X72736; CAAS1272.1; JOINED.
EMBL; X72737; CAAS1272.1; JOINED.
EMBL; X72738; CAAS1272.1; JOINED.
EMBL; X72739; CAAS1272.1; JOINED.
EMBL; X72740; CAAS1272.1; JOINED.
EMBL; X72741; CAAS1272.1; JOINED.
EMBL; X72742; CAAS1272.1; JOINED.
EMBL; X72743; CAAS1272.1; JOINED.
EMBL; S57777; AAB19988.1; -
PIR; S35990; KJHUAB.
PDB; 1FSU; 04-FEB-98.
Genew; HGNC:714; ARSB.
MIM; 253200; -
MIM; 272200; -
GO; GO:0005764; C-lysosome; TAS.
GO; GO:0004065; Fatty sulfatase activity; TAS.
GO; GO:0007041; P-lysosomal transport; TAS.
InterPro; IPR00917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE 1; 1.
PROSITE; PS00149; SULFATASE 2; 1.
Hydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharidosis;
Disease mutation; Polymorphism; 3D-structure.
SIGNAL 1 36
CHAIN 37 533
ACT SITE 147 147
MOD RES 91 91
DISULFID 117 521
DISULFID 121 155
DISULFID 181 192
DISULFID 405 447
CARBOHYD 188 188
CARBOHYD 279 279
CARBOHYD 291 291
CARBOHYD 366 366
CARBOHYD 426 426
CARBOHYD 458 458
VARIANT 92 92
VARIANT 95 95
VARIANT 117 117
VARIANT 137 137
VARIANT 152 152
VARIANT 160 160
VARIANT 210 210
VARIANT 236 236
VARIANT 302 302
VARIANT 376 376
VARIANT 393 393
VARIANT 405 405
VARIANT 498 498
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .) (PROBABLE).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .) (POTENTIAL).
T -> M (IN MPS-VI; MILD FORM).
/FTID=VAR 007294.
R -> Q (IN MPS-VI; MILD/SEVERE FORM).
/FTID=VAR 007295.
C -> R (IN MPS-VI; SEVERE FORM).
/FTID=VAR 007296
G -> V (IN MPS-VI; INTERMEDIATE FORM).
/FTID=VAR 007297.
R -> W (IN MPS-VI; INTERMEDIATE FORM).
/FTID=VAR 007298.
R -> Q (IN MPS-VI; INTERMEDIATE FORM).
/FTID=VAR 007299.
Y -> C (IN MPS-VI; MILD/INTERMEDIATE).
/FTID=VAR 007300.
L -> P (IN MPS-VI; MILD FORM).
/FTID=VAR 007301.
G -> R (IN MPS-VI; SEVERE FORM).
/FTID=VAR 007302.
V -> M.
/FTID=VAR 007303.
H -> P (IN MPS-VI; MILD/SEVERE FORM).
/FTID=VAR 007304.
C -> Y (IN MPS-VI; MILD FORM).
/FTID=VAR 007305.
L -> P (IN MPS-VI; MILD/SEVERE FORM).

FT CONFLICT 358 358 /FTID=VAR 007306.
FT STRAND 46 52 V -> M (IN REF. 3).
FT TURN 57 58
FT HELIX 61 63
FT TURN 64 64
FT HELIX 70 77
FT TURN 78 78
FT STRAND 80 82
FT TURN 83 84
FT STRAND 85 86
FT HELIX 93 100
FT HELIX 104 107

Query Match 4.4%; Score 210.5; DB 1; Length 533;
Best Local Similarity 20.8%; Pred. No. 1.2e-07;
Matches 102; Conservative 82; Mismatches 173; Indels 133; Gaps 23;

Qy 2 GPESLVLCCLLSATVFSLL---GGSSAFLSHRLKGRFQDRNRNIRENIIILVLTDD---QD 55
Db 14 GPRRLLPVVLPLLLLLAPPQSGAGSR-----PHELVLADLGLWND 59
Qy 56 VELGSMQVWVKTERIMEQGAHFINAFVTPMCCPSRSSILTQY-----VHNHNTYNN 110
Db 60 VGFHSGKIRTPHLDALAGGV-LLDNYYTOPLCTPSRSOLLTGRYQIRGTGLQHIWPCQ 118
Qy 111 ENCSSPSWQAOHESRTFAYVILNSTGYRTAFPGK-YLNEYNGSVYPP--GWKEWVGLLKN 167
Db 119 PSC-----VPLDEKLLPOLLEKAGYTTMVGKWHLMYRKCLPFRGFDTYFGVLLGS 172
Qy 168 RFY-----NYTLC---RNGVKEKHSYSDYLDLTDLITNDSVSFFRTSKMY 211
Db 173 EDYYSHERCTLIDALNVTRCALDFRDG--EEVATGYKNWYSNIFTKRALITNHP--- 227
Qy 212 PHRPVLWVISHAAPHGPDGSAQYSLFPNASCHITPSSYVAPNPKHMYGTGPKFI 271
Db 228 PEKFLFLYLALQSVHEPLQVPEYLK-----PYDFIQDKNRH---HYAG----- 268
Qy 272 HMEFTNMLQKRLQTLMSVDDSMETIYNNMLVETGELDNIYIVYTADHGYHIGQFGLVKGK 331
Db 269 -----MVSLM--DEAVGNTAALKSSGLWNTVFIETDNG---QQ-TLAGGN 310
Qy 332 SMP-----YEFDIRVPFYVGRPNV-EAGCLNPHIVLNIDLAFTILDIA--GLDIPAD 380
Db 311 NWPLRGRKWSLWEGGVGVGFVASFLLKQKGVKNRELIHISDWLPTLVKLARGHTNGTKE 370
Qy 381 MDGKSLIKLIDTERPVNRFHLKCKMVRWEDSFIVERGKLLHKEDNDKVD----- 429
Db 371 LDGFDVWKTISEGSPSPRI-----ELRNIDPNFVDSPPCPNSMAP 412
Qy 430 AQEENFLPKY 439
Db 413 AKDDSSILPEY 422

Search completed: February 14, 2004, 22:54:56
Job time : 44 secs